NBER WORKING PAPER SERIES

DOES LEED CERTIFICATION SAVE ENERGY? EVIDENCE FROM RETROFITTED FEDERAL BUILDINGS

Karen Clay
Edson R. Severnini
Xiaochen Sun

Working Paper 28612
http://www.nber.org/papers/w28612

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021, Revised August 2023

We are grateful for invaluable comments and suggestions from two anonymous referees, Andrea La Nauze, Josh Linn, Vivian Loftness, Kathleen Segerson, Yujie Xu, seminar participants at Carnegie Mellon University, and conference participants at the Association for Mentoring and Inclusion in Economics (AMIE)'s Workshop in Applied Microeconomics. We also gratefully acknowledge financial support from the Heinz College at Carnegie Mellon University. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Does LEED Certification Save Energy? Evidence from Retrofitted Federal Buildings
Karen Clay, Edson R. Severnini, and Xiaochen Sun
NBER Working Paper No. 28612
March 2021, Revised August 2023
JEL No. O31,Q41,Q48,Q52

ABSTRACT

This paper examines the causal impact of LEED (Leadership in Energy & Environmental Design) certification on energy consumption among federally owned buildings that were retrofitted over the period 1990-2019. Using a difference-in-differences propensity score matching approach, the paper has two findings. First, despite energy savings being an explicit federal goal, LEED-certified retrofits of federal buildings did not have statistically significant energy savings on average. Second, LEED buildings with higher energy scores had greater energy efficiency post-certification, and the improvements were economically meaningful. The absence of energy savings on average appears to be driven by three factors—trade-offs across energy and other areas in acquiring points for certification, possible changes in energy use after the official performance period for LEED certification ended, and improvements in the energy efficiency of all federal buildings.

Karen Clay
Heinz College
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
and NBER
kclay@andrew.cmu.edu

Edson R. Severnini
H. John Heinz III College
Carnegie Mellon University
4800 Forbes Ave #2114B
Pittsburgh, PA 15213
and IZA
and also NBER
ersevernini@gmail.com

Xiaochen Sun
4800 Forbes Avenue
Pittsburgh, PA 15213
xiaoches@andrew.cmu.edu

1 Introduction

Federal building energy efficiency efforts began in the 1990s and gained momentum in the 2000s. The Energy Independence and Security Act was signed in 2007 "to improve the energy performance of the Federal Government" (F.R., 2007), and defines high-performance building as "a building that integrates and optimizes on a life cycle basis all major performance attributes, including *energy conservation*." In response to these regulations, the U.S. General Services Administration (GSA) announced a $5.55 billion plan in 2010 to achieve significant energy savings from their buildings. To implement this, the federal government decided to use two third-party green building certification systems, Green Globes and LEED (Leadership in Energy & Environmental Design), and not the EPA's Energy Star program (U.S. GSA, 2013). Although Green Globes certification was permissible, almost all of the actual certifications were LEED certifications.

Whether LEED would generate actual energy savings in federal buildings is not a priori clear. There might be selection bias in the buildings that are retrofitted. Even controlling for selection bias, three other factors may play roles. First, decision makers face budget constraints and so may choose to trade-off energy efficiency and non-energy efficiency dimensions to acquire the points necessary for LEED certification. Second, there could be behavioral changes from occupants and building after certification. Studies have shown that behavioral responses such as increasing heating or cooling or sub-optimal maintenance can cut down the intended energy savings by more than 5% (Gillingham et al., 2013). Third, there could be improvements in energy efficiency happening in non-LEED federal buildings.

In this study, we examine the causal impact of LEED certification on energy consumption among federally owned buildings that were retrofitted over the period 1990-2019.[1] It is worth noting that one common challenge in the green building literature is that many certified buildings are new and data points become scarce when restricting the consumption data to a longer period of time (Matisoff, Noonan and Flowers, 2016).[2] Our paper draws on energy use data for office buildings and courthouses owned by the federal government across the entire U.S. Specifically, we combine energy use data from the GSA's

[1]It is important to point out that *existing buildings* can obtain LEED certification as *new construction* if the retrofits exceed 60% of the building square footage.

[2]Notably, the federal government published a contracted third-party report that drew conclusions on only 22 buildings (Fowler et al., 2010).

Energy Usage Analysis System with data from the Green Building Information Gateway and weather data from National Oceanic and Atmospheric Administration. Buildings in our sample represent 84% of the square footage of federally owned LEED certified existing building space.

Following Eichholtz, Kok and Quigley (2010, 2013), we use a difference-in-differences propensity score matching approach to estimate the causal effect of LEED certification on energy efficiency. Matching methods tackle selection on observables, and difference in differences controls for the presence of unobservables that under normal circumstances may lead to biased estimates. These methods address two issues – possible selection bias in buildings that are retrofitted and the fact that LEED certified buildings tend to be much larger than average. The difference-in-differences estimation compares buildings that were LEED certified before and after certification with otherwise similar buildings that were not LEED certified. We have a large never treated group, so staggered difference-in-differences estimates are unlikely to be biased (Marcus and Sant'Anna, 2021). To verify this, we provide supporting evidence from a Goodman-Bacon (2021) decomposition.

The paper has two main findings. First, despite energy savings being an explicit federal goal, LEED-certified retrofits of federal buildings did not have a statistically significant effect on energy savings on average. The point estimate for the causal impact of LEED certification on energy consumption is -3.0%, and we cannot rule out effects from -12.0% to 6.0%. The evidence is consistent with three factors contributing to a lack of savings. As we describe further below, we find evidence consistent with decision makers making trade-offs in the face of budget constraints. The event study suggests there may have been changes in energy use after the official performance period for LEED certification ended. The data shows there were improvements in the energy efficiency of all federal buildings.

Second, we find that LEED buildings with higher energy scores had statistically significantly greater energy efficiency post-certification. Having a one standard deviation higher energy score is associated with 12.6% lower energy usage in all buildings and 13.9% lower usage in office buildings. Some other attributes, notably higher water scores, are associated with lower energy efficiency post-certification. The decision makers involved with retrofitting federal buildings face budget constraints, so greater expenditure on water reduction may lead to lower expenditure along other dimensions that impact energy usage. Our limited data on water use intensity is consistent with this trade-off.[3]

The main results are robust to a wide variety of sample and specification checks. We show they are robust to controlling for the Energy Star Status of treatment and control buildings. We explore alternative matching conditions such as defining pre-treatment usage based on a common calendar time to account for common time-varying shocks, remove repeated observations due to matching with replacement and run unweighted regressions, compare the estimates arising from the matched sample and original sample, drop the retrofits deemed new construction, consider specifications in levels rather than in logs, and conduct analysis on a sample including up to six instead of four years after certification, among others. Importantly, in a decomposition proposed by Goodman-Bacon (2021), we find that more than 91 percent of the variation used to obtain the overall difference-in-differences estimates comes from the comparison between treated and never treated buildings, with no negative weights.[4] Therefore, it is unlikely that our estimates are biased due to the variation in treatment timing.

Our paper contributes to two literatures. The first is the literature on energy savings in LEED buildings.[5] The U.S. Green Building Council states that there were $1.2 billion in energy savings from the LEED certified buildings between 2015 and 2018 (USGBC, 2020*b*). These estimates, however, are based on comparisons with guidelines rather than examining the actual performance of buildings. A number of engineering studies have compared the energy efficiency of LEED and non-LEED buildings.[6] Although they found no energy savings, the analysis was not causal. In an appendix analysis using a methodology similar to ours, Qiu and Kahn (2019) estimated the impacts of LEED certification for commercial buildings in Phoenix, Arizona, and found a zero effect as well. Data from private commercial buildings are proprietary and, therefore, difficult to access. Our pa-

[3]Water reduction technology in some cases may also consume more energy or be paired with more energy intensive technologies. For example, this could be the case if the electricity consumption of sensor technology in bathrooms or for landscaping is relatively high. Or it could occur if sensor enabled bathrooms tend to include electric hand driers to replace paper towels.

[4]As discussed in Goodman-Bacon (2021), negative weights only arise when average treatment effects vary over time. Our event study analysis does not point to that case. In fact, over more than half of the LEED-certified retrofitted buildings in our sample were registered in one single year – 2012.

[5]In a related literature on the market effect of LEED certification, Eichholtz, Kok and Quigley (2010, 2013) find that LEED-certified commercial buildings command substantially higher rents and selling prices than otherwise identical buildings. Their evidence suggests that the "intangible" effects of the label itself may play a role in determining the values of green buildings in the marketplace. In fact, Matisoff, Noonan and Mazzolini (2014) find that marketing benefits due to LEED certification may be the primary drivers in the development of green buildings.

[6]See Wedding and Crawford-Brown (2008); Scofield (2009, 2013); Scofield and Doane (2018).

per builds on this literature by providing the first nationwide causal estimates of energy savings from the LEED program using data from a national sample of federal government buildings.

The second is the energy gap literature. This literature documents the failure of energy improvements to yield expected outcomes (e.g., Allcott and Greenstone, 2012; Levinson, 2016; Gerarden, Newell and Stavins, 2017; Fowlie, Greenstone and Wolfram, 2018; Gillingham, Keyes and Palmer, 2018). Most of the analysis has been in the residential sector. Exceptions are Kahn, Kok and Quigley (2014) and Brolinson, Palmer and Walls (2023), although neither of their sample includes government buildings. Kahn, Kok and Quigley (2014) find that commercial buildings that are newer and of higher quality consume more electricity. The literature identifies a number of reasons for energy gaps such as inattention and myopia, the rebound effect, and quality of installation and worker incentives.[7] Brolinson, Palmer and Walls (2023) show no energy savings associated with Energy Star certification because certified buildings are already energy efficient beforehand. Our paper builds on this literature by documenting three possible reasons for the failure of LEED certification to generate energy savings in federal buildings.

In what follows, Section 2 lays out the policy background and introduces the LEED program. Section 3 describes the data and summary statistics of the buildings in the analysis. Section 4 presents the empirical strategy. Section 5 reports the (lack of) energy savings from the LEED program, and discusses the heterogeneity of the treatment effects identifying a mechanism behind the estimated energy outcome. Section 6 offers some concluding remarks.

2 Background

The building sector represents 39% of the total energy consumption in the U.S. (U.S. EIA, 2021). A substantial portion of that consumption is thought to be wasted (U.S. DOE, 2020). Federal government buildings are regarded as less efficient and consume considerably more energy than non-government commercial buildings (U.S. EIA, 2016).

[7]For reviews of this literature, see Gillingham, Newell and Palmer (2006, 2009), Allcott and Greenstone (2012), Gerarden, Newell and Stavins (2017), Gillingham, Keyes and Palmer (2018), and Myers (2019). See also Giraudet, Houde and Maher (2018); Blonz (forthcoming); Christensen et al. (2023) for moral hazard issues related to energy efficiency.

As a result, there has been a growing attention to green building design and an increasing number of policies that regulate energy performance in federal buildings.

Energy efficiency efforts began in the 1990s and gained momentum in the 2000s. The standards for energy efficiency for federal buildings were first set by the Energy Policy Act in 1992 (Sharp, 1992). It comprehensively described the energy needs in the United States and aligned the building codes with energy efficiency codes. This policy was amended in 2005 to set energy reduction goals for the federal buildings between 2006 and 2015 (Pombo and Thomas, 2005).

The Energy Independence and Security Act (EISA) was signed in 2007 "to improve the energy performance of the Federal Government" (F.R., 2007), and defines high-performance building as "a building that integrates and optimizes on a life cycle basis all major performance attributes, including *energy conservation*." Along with EISA, Executive Orders 13423, 13514, 13636, and 13834 required that "[t]o improve environmental performance and Federal sustainability, *priority should first be placed on reducing energy use and cost*" (F.R., 2015, p.15871, our emphasis).

In response to these regulations, the U.S. General Services Administration (GSA) announced a $5.55 billion plan in 2010 to achieve significant savings from their buildings. To implement this, the federal government decided to use third party green building certification systems, Green Globes and LEED, and not the EPA's Energy Star program. The reason stated was "saves resources by eliminating the cost to the government of developing its own duplicative green building certification system while drawing on the expertise of the private sector" (U.S. GSA, 2013, p.1).

Although Green Globes certification was permissible, almost all of the actual certifications were LEED certifications. LEED was developed by the U.S. Green Building Council (USGBC) to rate green building strategies across several categories. Unlike other widely recognized labels such as the Energy Star, which focuses on energy use, LEED certifies a building based on a comprehensive evaluation of six categories. The scorecard includes energy and atmosphere, materials and resources, indoor environmental quality, sustainable sites, water efficiency, and innovation in operations. The energy component accounts for the largest share, more than 30%. This paper does not focus on the heterogeneous effects across LEED tiers because we have a limited number of observations within each tier bin.[8] Instead, we investigate the impacts associated with the component scores directly.

Table 1 illustrates some important credit categories for each component in the LEED scorecard. For energy, buildings can get points for high efficiency lighting, HVAC, and insulation, as well as through the use of green power and energy audits. For water, buildings can get points by using water saving sink faucets, automated flushing, and high efficiency irrigation technology. For sustainable sites, buildings can earn points for alternative commuting. Other categories also offer a range of opportunities to earn points.

3 Data

This study uses five sources of data: the Energy Usage Analysis System (EUAS) dataset from the General Services Administration (GSA); the Green Building Information Gateway (GBIG) reports; the DOE Compliance Tracking System (CTS); the Registry of Energy Star Certified Buildings and Plants; and weather data from the National Oceanic and Atmospheric Administration (NOAA).

The EUAS dataset provides monthly energy consumption data for GSA buildings in the 48 contiguous American states plus District of Columbia for the period 1990-2019. We aggregate the data to the annual level, because the treatment variable is at the annual level. Building energy consumption can be measured as site energy or source energy. Site energy, which is reported by the EUAS dataset, is the amount of heat and electricity consumed by a building as reflected in the utility bill. Source energy represents the total amount of raw fuel that is required to operate a building, incorporating all transmission, delivery, and production losses (U.S. EPA, 2020).[9] Our results are similar across the two measures, so we report results for site energy in the main text and selected results for source energy in the appendix.

The GBIG reports provide information on LEED certified buildings. Based on the

[8]LEED-certified projects earn one of four rating levels: certified, silver, gold or platinum. Appendix I provides a more detailed introduction to the LEED program. Prior studies in real estate and marketing have examined tier bunching (Matisoff, Noonan and Mazzolini, 2014; Rysman, Simcoe and Wang, 2020), but we lack enough data variation in our setting.

[9]Since both site and source energy can be served as the benchmark for LEED compliance (USGBC, 2018), we calculate source energy using the U.S. EPA (2019) Technical Reference.

street address, we match these reports to the EUAS dataset. Specifically, it includes facility address, program registration and certification dates, certified square footage, and scores for each LEED category (see Table 1 for examples of score categories). The NOAA data allow us to control for variation in weather across time and building location.

CTS has information on upgrades that occurred in both LEED buildings and non-LEED buildings. A third of the buildings in the dataset have adopted efficiency conservation measures (ECM), regardless of their LEED status. These upgrades include advanced metering, building envelope, lighting improvements, updated HVAC, and water and sewer conservation measures. Buildings might adopt more than one of these upgrades, and they are not necessarily completed in the same year. Appendix Figure A.1 shows the completion year of the conservation initiatives the buildings have undertaken. Most conservation measures were finished in 2009 and 2016 for both LEED and non-LEED buildings. There is no indication that LEED certification requires those upgrades. In fact, the correlation between LEED certification and participation in any energy efficiency initiative is remarkably low (-0.062). Furthermore, some LEED buildings completed a conservation program before registration, while others finished the upgrades after registration (Appendix Figure A.2).

The Registry of Energy Star Certified Buildings and Plants dataset provides information on Energy Star labeling in the GSA buildings. Approximately 30% of the non-LEED buildings and 40% of the LEED buildings in the dataset were labeled Energy Star. Most of the labels were awarded in 2000, 2008, and 2009. The correlation between LEED certification and Energy Star labeling is also very low (0.042).

Our analysis focuses on 60 LEED-certified retrofitted federally-owned buildings with complete LEED component score data. The EUAS dataset provides energy consumption data for 110 LEED and 452 non-LEED GSA buildings. We restrict attention to buildings owned rather than leased by the federal government, because the LEED buildings are all federally owned. We keep only buildings with at least 4 years of data before and 4 years of data after the treatment date to ensure we have enough data for the pre-post comparison in the difference-in-differences estimation. This drops 43 LEED buildings, many of which are new construction. A few LEED buildings are dropped due to incomplete data.[10] See Appendix Table A.1 for information on the number of buildings in the original dataset

[10]GSA also endorsed Green Globes as an alternative third-party green building certification system, but only *five* GSA buildings were Green Globes certified in our dataset.

as well as the reductions associated with each sample restriction.

Four issues are worth noting with respect to building data generally and LEED buildings specifically. First, building property data is limited, especially for uncertified buildings. The most commonly used dataset, CoStar commercial real estate data, only includes high-quality commercial office space, which raises questions about the external validity of the estimates. In addition, consumption data is also confidential for most buildings, making it hard to evaluate building performance. Notably, the federal government published a contracted third-party report that drew conclusions on only 22 buildings (Fowler et al., 2010). Our paper, on the other hand, leveraged publicly available datasets to estimate the outcome of LEED with *all* federal buildings in the U.S., in order to strike a balance between internal validity and external validity.

Second, GBIG reports provide both registration and certification dates for LEED buildings. Registration indicates intent to pursue LEED certification and is often done in the design phase. Certification occurs after the retrofit is complete, collection of any performance data has occurred, and all paperwork has been submitted (Wehe, 2020). Figure 1 shows the distribution of registration and certification years. The modal year for registration is 2012 and for certification is 2013. We analyze the effects of LEED using both the registration and the certification dates as the treatment date. As we see in section 5.3, the two dates give similar results. Because it avoids classifying potentially treated years as untreated, our main focus is on the registration date.[11]

Third, buildings can be LEED certified for a subset of the overall square footage. Figure 2 shows the distribution of the certified ratio. Most LEED-certified buildings in our sample obtain certification for a high share of the overall square footage. The guidelines published by the USGBC (2020*a*) use 5% as the lower bound for certification.

Fourth, *existing buildings* can obtain LEED certification as *new construction* if the retrofits exceed 60% of the building square footage. Appendix Table A.2 shows the type of LEED programs for the 60 LEED-certified retrofitted buildings used in our main analysis. More than half of those buildings received certification from LEED Existing Building programs, and a third received certification from LEED New Construction programs.

[11] This is analogous to McCrary (2007)'s definition of treatment in another context. He was interested in estimating of the effect of court-ordered affirmative action in police hiring on workforce composition, but uses the onset of litigation rather than the court ruling date as the beginning of treatment.

Table 2, Panel A, provides summary statistics for LEED and non-LEED buildings, including pre-treatment energy use intensity and building size. Energy use intensity (EUI) is defined as the ratio of total energy consumption and the square footage of a building. It also reports the differences between treated and untreated building groups. LEED buildings are significantly larger than non-LEED buildings in square footage. The federal government appears to have focused on the largest buildings for LEED certification. The majority of LEED and non-LEED buildings are office buildings, and most of the remaining buildings are courthouses.

Table 2, Panel B, provides the number of LEED and non-LEED buildings by building type, and Panel C reports average LEED program attributes in our sample. Because the LEED program has changed slightly over time, we report the component score divided by the total achievable score. Given that the component scores for materials and resources, indoor environmental quality, and innovation in operations are highly positively correlated (see Appendix Table A.3), we combined these into a single measure. Histograms of the scores are shown in Appendix Figure A.3. To facilitate comparison, in the analysis the component scores and the certification ratio are all standardized to have mean zero and standard deviation one.

Figure 3 shows there were improvements in the energy efficiency of all federal buildings over the sample period. Average energy use is similar for LEED and non-LEED buildings throughout the period of our analysis. Energy consumption started declining for both groups around 2010, and LEED buildings do not appear to save more energy compared to non-LEED buildings. The improvement in energy efficiency in federal buildings starting 2010 likely reflects a series of policies during the Obama Administration, including the Executive Order 13514 signed in 2009 and the Better Building Initiative launched in 2011 (Obama, 2011). These initiatives required the federal agencies to achieve zero net energy by 2030.

Table 3 shows descriptive regressions of LEED attributes on energy use intensity in the year when LEED buildings were certified. We find that the energy component score is negatively correlated with energy consumption while the water component score is positively correlated with the consumption during the certification year.

4 Empirical Strategy

Building on Eichholtz, Kok and Quigley (2010, 2013), we implement a two-stage strategy to address differences in pre-treatment characteristics of the LEED and non-LEED buildings, and then compare the two groups to infer the energy reductions brought about by the LEED program. Specifically, propensity score matching is employed in the first stage to construct a sub-sample of the untreated buildings with similar characteristics to treated buildings. Difference-in-differences models are applied in the second stage with the newly constructed sample to evaluate the energy outcome from registration for certification.

Our paper implements several matching algorithms and estimates both site and source energy savings from the LEED program. These specifications generate similar estimates. Thus, we report the results for site energy with a five nearest-neighbor (5NN) propensity score matching in the main discussion, leaving the three nearest-neighbor (3NN) and source energy cases for the robustness check section and appendix.

4.1 First Stage: Propensity Score Matching

In the first stage of the analysis, we regress the LEED indicator on building characteristics and obtain the propensity scores. Following Caliendo and Kopeinig (2008), we use a *probit* regression to bound the LEED indicator in the range of $[0, 1]$.

The basic specification is

$$Pr(D_i = 1|X_i) = \Phi(X_i'\gamma), \tag{1}$$

where D is a dummy variable equal to one if a building i is LEED certified and zero otherwise. X represents an array of building characteristics that may differ between the treated and untreated buildings, including building size and the pre-treatment energy use intensity (EUI). Moreover, we also incorporate the polynomial forms of these variables in the X to address biases that might arise from a specific functional form.[12]

[12] The estimated propensity scores become stable after having the energy consumption and building size terms with third or higher power. For this reason, the X includes: building size^n ($n = 1, 2, 3$) and pre-treatment site $log(EUI)^n$ ($n = 1, 2, 3$).

Figure 4a shows the distribution of the propensity scores before matching. The horizontal axis represents the propensity score while the vertical axis corresponds to the proportions of buildings that fall into each propensity score bin. It indicates good common support if the distribution of the propensity scores is similar for the two groups. We see from the figure that the treated buildings are evenly distributed across the propensity score range, while the untreated buildings are right-skewed with many of them falling in the first propensity score bin. With the estimated propensity score, we match the non-LEED buildings to the LEED buildings using k-nearest neighbors (kNN) with replacement. This method matches k untreated buildings to each treated building that share similar propensity scores, so as to create a subset of buildings that are comparable in size and baseline energy consumption.[13]

Figure 4b shows the propensity score distribution for treated and untreated buildings after matching with five nearest neighbors (5NN). We see that the untreated buildings are now more evenly spread across the propensity score bins. More importantly, the distributions of the scores are similar across the two groups, which supports the validity of the matching results. In this matching process, every building in the treatment group is able to find a relatively good match. Figure 5 shows the locations of LEED and non-LEED buildings before and after matching.

Table 2, Panel A, also reports the descriptive statistics of LEED and non-LEED buildings after matching with 5NN. The gap in building size between the two groups is greatly reduced after matching and is no longer statistically significant. The difference in the pre-treatment energy use intensity remains insignificant. These statistics provide evidence for the common support assumption in the second-stage analysis.

4.2 Second Stage: Difference-in-Differences Models

In the second stage of the analysis, we estimate the energy savings from the LEED program. The specification is as follows:

$$ln(EUI_{it}) = \theta T_{it} + X_{it}'\beta + \gamma_i + \lambda_t + \epsilon_{it}, \tag{2}$$

[13] If the propensity score of a treated building is higher than the maximum or lower than the minimum score of any untreated buildings, this treated unit would be considered an outlier and dropped from the dataset. None of the LEED buildings are dropped.

where subscript i refers to each individual building, and t refers to each year. The dependent variable $ln(EUI_{it})$ is the energy use intensity in logarithm form. T_{it} is the treatment indicator equal to 1 if building i in year t is registered for LEED certification. X_{it} is a vector that captures weather variation, which includes annual cooling degree days (CDDs) and annual heating degree days (HDDs). γ_i represents a set of building fixed effects, and λ_t a set of year fixed effects. Because the matching with replacement in the first stage can introduce duplicate buildings in the untreated group, the regression is weighted by the *inverse* of number of repeated observations in the dataset.[14]

To examine the parallel trends assumption, we carry out an event study analysis to show the energy consumption difference between the treated and untreated groups each year relative to the difference in the reference year. Because the LEED buildings are registered in different years, we treat the registration year as year zero. After that, lead and lag terms are created for years before and after the registered year. Our specification leaves out the -1 year from the regression, which is the year immediately prior to registration. Specifically, the regression is run as follows

$$ln(EUI_{it}) = \gamma_i + \lambda_t + \sum_{\tau=2}^{q} \delta_{+\tau} D_{i,t+\tau} + \sum_{\tau=0}^{m} \delta_{-\tau} D_{i,t-\tau} + \tilde{D}'_{it}\delta + X'_{it}\alpha + \epsilon_{it}. \tag{3}$$

The first sum allows for q leads (δ_{+1},...,δ_{+q}) and $D_{i,t+\tau}$ equals to 1 if the calendar year t is τ year(s) before the building is registered for LEED certification. Similarly, the second sum allows for m lags (δ_0, ...,δ_{-m}) and $D_{i,t-\tau}$ equals to 1 if the calendar year t is τ year(s) after LEED certification. Because our sample includes at least 4 years pre- and post-treatment, q and m are both set to be 4. $\tilde{D}'_{it}$ is a vector that captures the sum of lead and lag terms that are far away from the treatment year to minimize noise in the estimation, and to reduce multicollinearity with the set of building and year fixed effects. It represents the endpoint restrictions, as explained by Kline (2012).

Although the treatment timing varies – buildings may register in different years – identification can rely on the weakest parallel trends assumption considered by Marcus and Sant'Anna (2021) because we have a never treated group. Such assumption does not impose any restriction on pretreatment trends across groups. In fact, when the number

[14]More than half of the non-LEED buildings are matched only once, but there are cases when non-LEED buildings are matched more than once. The maximum number of matches is 7. In the robustness check section and appendix, we also present results weighted by both building size and duplicates.

of never treated units is "reasonably large" – as it is the case here – that assumption can identify policy-relevant parameters even "if researchers are not comfortable with a priori ruling out nonparallel pretrends" (Marcus and Sant'Anna, 2021, p.251).[15]

Since there is variation in treatment timing, we implement the decomposition proposed by Goodman-Bacon (2021). To preview our results, we find that more than 91 percent of the variation used to obtain the overall difference-in-differences estimates comes from the comparison between treated and never treated buildings, with no negative weights.[16] This is not surprising because more than half of the LEED-certified buildings in our sample are registered in one single year. Therefore, it is unlikely that our estimates are biased due to the staggered nature of the treatment.

5 Results

5.1 Average Impacts on Energy Consumption

Table 4 shows that LEED-certified retrofits of federal buildings did not lead to statistically significant energy savings on average.[17] In column 1, the point estimate for the causal impact of LEED certification on energy consumption is -3.0%, and we cannot rule out effects from -12.0% to 6.0%.[18] Columns 2 and 3 include treatment interacted with LEED component scores and the certification ratio and examine the effects separately for all buildings and for office buildings. The average effects of LEED-certified retrofits remain small and not statistically significant.[19] Columns 4-7 add indicator variables for whether a building had Energy Start certification and whether it implemented energy conservation measures. The results still remain small and not statistically significant. These results

[15]On the other hand, de Chaisemartin and D'Haultfœuille (2020) assume a stronger parallel trends assumption that restricts all pretrends in all pretreatment periods. That is, in the absence of treatment, the expectation of the outcome of interest follows the same path in all groups and in all time periods available in the data.

[16]As discussed in Goodman-Bacon (2021), negative weights only arise when average treatment effects vary over time. Our event study analysis does not point to that case.

[17]Appendix Figure A.4 shows the full distribution of energy consumption for LEED and non-LEED buildings, before and after registration.

[18]Abadie (2020) argues that "rejection of a point null often carries very little information, while failure to reject may be highly informative" (p.193). He "challenge[s] the usual practice of conferring point null rejections a higher level of scientific significance than non-rejections," and "advocate[s] visible reporting and discussion of nonsignificant results" (p.193).

[19]The certification ratio and the component scores are all standardized to have a mean of zero and a standard deviation of one.

are consistent with the patterns in Figure 3 and with increasing emphasis on energy savings in all federal buildings.

Figure 6a presents the results of the event study showing the effects on energy consumption. The x-axis represents a timeline of the LEED certification procedure. Buildings are registered at year 0, and the reference year is set at year -1. The average certification year of the LEED buildings is 3 years after the registration date. The y-axis is the energy savings from the LEED program in each year relative to the reference year. The pre-trend looks flat and seems to supports the parallel trends assumption needed for our difference-in-differences approach.

In the event study, energy use increased from the year prior to certification to the first certified year, hinting that behavioral or other changes may have played a role in the lack of energy savings.[20] In most cases the year prior to certification is a performance period where data is collected for certification. This increase may reflect increased use of heating or cooling, increased use of hot water, or poor maintenance of energy-related building systems (Gillingham et al., 2013).

5.2 Heterogeneous Effects by LEED Scores

In Table 4, LEED buildings with higher energy scores had statistically significantly greater energy efficiency post-certification. In columns 2 and 3, having a one standard deviation higher energy score is associated with 12.6% lower energy usage in all buildings and 13.9% lower usage in office buildings, respectively. Energy Star and energy conservation measures (ECM) are included as control variables in columns 4-7. Although both Energy Star and ECM seem to reduce energy consumption, we continue to find that LEED certification did not generate energy savings. We explore the effects of Energy Star further in Appendix Table A.4 by interacting Energy Star with LEED certification and subcategory scores and by dropping Energy Star buildings. The main effect of LEED certification on energy use intensity remains small and statistically insignificant. In Table 4, buildings with higher certification ratios had statistically significantly lower energy consumption. As we discuss further below, buildings with higher water scores had

[20] It is worth noting that the observed increase is statistically insignificant. We conduct several joint F-tests for the coefficients of these event periods and find them to be jointly insignificant. The p-values for the joint coefficients are 0.7162 for periods 0, 1, and 2; 0.8933 for periods 3 and 4; and 0.8996 for all time periods.

statistically significantly higher energy consumption.

The effect of a one standard deviation higher energy score is economically meaningful. The average energy intensity of the LEED buildings in the year they were certified is approximately 18kWh/ft^2. For a 400,000 square foot building, the savings would be approximately 907,200 kWh per year. An average home in the United States consumes 11,000 kWh per year, so the savings would be equivalent to consumption by 82 homes. The average electricity price nationally in 2020 was 11 cents/kWh, the annual savings associated with the reduction would be 0.25/ft^2 or about $100,000 for a 400,000 square foot building.[21] The typical cost of green retrofit for an entire building is anywhere between $2-7/ft^2 (Lockwood, 2009), and thus the average payback year of the LEED retrofit based only on energy savings is around 8-28 years. This implied payback period is reasonable, given that the usual payback time for an energy efficiency program is around 2-15 years depending on the building's age and other information (Lockwood, 2009). Moreover, the payback period is relatively consistent with the requirements from the Obama (2011)'s Memorandum if the retrofit is done properly. The Memorandum specified that "agencies shall fully implement energy conservation measures (ECMs) in federal buildings with a payback time of less than 10 years."

Table 4 also shows that LEED buildings with higher water scores had lower energy efficiency post-certification. In columns 2 and 3, having a one standard deviation higher water score is associated with 6.3% higher energy usage in all buildings and 8.3% higher usage in office buildings, respectively. The results are similar in columns 4-7. [22]

The adverse effects of water scores on energy efficiency likely reflects budgetary trade-offs.[23] The decision makers involved with retrofitting federal buildings make decisions under those constraints, so greater expenditure on water reduction may lead to lower expenditure along other dimensions that impact energy usage. Although we have limited data on water use intensity, the results in Appendix Table A.6 are consistent with

[21] The electricity price varies across states. Based on the U.S. EIA (2020)'s monthly report, California has the highest electricity price, 18.39 cents/kWh, while Nevada has the lowest electricity price, 4.59 cents/kWh.

[22] Appendix Table A.5 calculates the effects for different combinations of the energy and water scores using coefficients from Table 4 columns 2 and 3.

[23] Water reduction technology in some cases may also consume more energy or be paired with more energy intensive technologies. For example, this could be the case if the electricity consumption of sensor technology in bathrooms or for landscaping is relatively high. Or it could occur if sensor enabled bathrooms tend to include electric hand driers to replace paper towels.

budgetary trade-offs.[24] The coefficient on water score is not statistically significant. The coefficient on energy score is positive and statistically significant. In columns 3 and 4, having a one standard deviation higher energy score is associated with 8.3% higher water usage intensity in all buildings and 9.3% higher usage in office buildings, respectively.

5.3 Robustness Checks

This section shows that our results are robust to a range of possible concerns about the analysis. Specifically, we examine the sensitivity of our results to the empirical specification, the matching procedure, and the sample.

5.3.1 Specification

One concern relates to the sensitivity of our results to the empirical specification. Table 5 shows in columns 1-4 that our results are robust to using the certification year rather than the registration year as the treatment year, and to estimating the effects in levels rather than logs.

We consider the potential bias from occupancy changes in Appendix Table A.7. [25] Appendix Table A.8 shows the space standards, criteria, and guidelines for federal buildings, where the number of employees per office space is explicitly regulated. Therefore, it is unlikely for a building to significantly increase its population density, which would lead to an increase in energy consumption.[26] According to a third-party report published by DOE, "... occupancy was a contributing factor to energy consumption but the contribution was not as strong as that of weather" (Selvacanabady and Judd, 2017). In Appendix Table A.7, we use CDD and HDD as proxies for weather changes and provide heterogeneity analysis regarding these two variables in column 2. The coefficients of the interaction terms show null effect of CDD and HDD on energy use intensity. Hence, if anything, we should expect an even smaller estimate for the occupancy level changes.

We use staggered difference in differences, which can lead to biased estimates in

[24]Water use intensity data are only available for about half of the sample.

[25]Note that our paper has already used energy consumption per square footage as the outcome variable in our main specification to implicitly control for potential variation introduced by building usage intensity.

[26]U.S. EIA (2022) reports that the number of workers per square foot is decreasing over time in *commercial* buildings.

some contexts. The Goodman-Bacon (2021) decomposition in Table 6 suggests that it is unlikely that our estimates are biased due to the variation in treatment timing. More than 91 percent of the variation used to obtain the overall difference-in-differences estimates comes from the usual comparison between treated and *never treated* buildings. Further, there are no negative weights. As mentioned before, this is not surprising given that we have a large untreated sample (Marcus and Sant'Anna, 2021) and that more than half of the LEED-certified buildings in our sample were registered in one single year (see Figure 1). Therefore, it is unlikely that our estimates are biased due to the variation in treatment timing.

As a falsification test, we also run the event study regressions only with the non-LEED buildings, artificially attributing them registration years. Although we implemented the propensity score matching, it would still be a concern if the untreated buildings are affected by some shocks that are missed from the difference-in-differences model but would drive down their energy use intensity. We repeat the first and second stage regressions with only the non-LEED buildings and report the estimated δ's from Equation (3). Detailed procedures for this falsification test can be found in Appendix III. If shocks do not play a major role in our findings, all the estimated δ's should be around zero. Figure 6b plots the results for 5NN matching. The coefficients are all close to zero. This pattern also holds when conducting the analysis with source energy and the 3NN matching, as displayed in Appendix Figure A.5.

5.3.2 Matching

Another concern is that our results are somehow sensitive to the specifics of the matching procedure. Table 5 shows in columns 5-10 that our results are robust to not matching at all and to matching by census region and census divisions. The robustness to not matching is not surprising, given that LEED and non-LEED buildings only differed in building size. To investigate the regional impacts on the energy component score and water component score in LEED buildings, Appendix Figure A.6 plots the score distributions across four census regions. Notice that the energy scores in the Midwest are comparable to those in the South, despite the contrasting climate zones of these two regions. Likewise, the water score distributions are identical across all four regions. Places with water scarcity problem, such as the West, do not have different water scores.

Appendix Table A.9 reports that our results are also robust to the number of nearest neighbors used in the matching procedure, and to using source rather than site energy.[27] For both site and source energy, the point estimates on the coefficients on energy score and water score are somewhat higher for 3NN than for 5NN matching. The estimates are not, however, statistically significantly different from our main estimates.

Appendix Table A.10 shows that our results are robust to a range of other matching strategies. Columns 1 and 2 consider an alternative matching strategy by defining pre-treatment usage based on a common calendar time to account for common time-varying shocks such as weather and macroeconomic conditions. Columns 3 and 4 drop the repeated non-LEED buildings in the untreated group due to matching with replacement. Columns 5 and 6 include the standard deviation of energy use intensity before treatment in the propensity score matching. Columns 7 and 8 match within building type. Although these restrictions tend to reduce the number of LEED buildings in the sample, the results follow the same pattern as our main findings reported in Table 4.

5.3.3 Sample

A third concern is that features of the sample are affecting the results. We explore this in two ways. First, Appendix Table A.11 columns 2-3 show the results with LEED buildings certified with and without New Construction. The estimates are qualitatively similar to our main results. The standard errors are relatively large due to sample size, so we cannot rule out that the estimates in each row are statistically of the same magnitude. Columns 4-5 separate the LEED buildings into buildings that were LEED certified within three years of registration (shorter time interval) and buildings that were LEED certified more than three years of registration (longer time interval).[28] Results are slightly different across the two columns, where the treatment indicator is significant when buildings were certified shortly after registration. But by large, the interactions point to the same direction and are similar to the main results. Appendix Table A.12 examines the effects of dropping buildings with low certification ratios, including observations with more than six months of data when annualized, requiring more post registration data, and weighting buildings both to account for matching with replacement and by square footage. The

[27] Appendix Figure A.7 suggests that the parallel trends assumption also seems to hold in these cases.

[28] Note that the average time interval between registration and certification date is three years for buildings in our dataset.

results in columns 1-8 are robust to these alternative samples.

Second, Appendix Table A.13 examines the effects of alternative time periods. Government efforts to improve energy efficiency in buildings become more prominent later in our sample period. Columns 1 and 2 duplicate our main results from columns 2 and 3 of Table 4, which include data for 1990-2019. Columns 3 and 4 shorten the time period to 1995-2019 and columns 5 and 6 shorten the time period to 2000-2019. The results in columns 3 and 4 are remarkably similar to our main results. In columns 5 and 6, the point estimates for the coefficients on energy score remain statistically significant but are somewhat smaller in magnitude. For example the coefficient is -0.084 versus -0.126 for all buildings. The point estimates for the coefficients on water score are smaller and no longer statistically significant. The coefficients on energy score and water score are not, however, statistically significantly different from the coefficients in columns 1 and 2. The results in columns 1-6 are robust to these alternative time periods.

6 Concluding Remarks

The U.S. federal government has implemented a series of policies to improve energy efficiency and sustainability of the federal buildings. LEED certification was one important focus of this effort. Using propensity score matching and difference in differences models, this paper investigated the causal impact of LEED certification on energy consumption in federally owned buildings that were retrofitted over the period 1990-2019.

The paper has two key findings. First, LEED-certified retrofits of federal buildings did not have statistically significant energy savings on average. This is despite energy savings being an explicit federal goal. Second, LEED buildings with higher energy scores had greater energy efficiency post-certification, and the improvements were economically meaningful. The available evidence points to three factors driving the lack of energy savings: trade-offs across energy and other areas in acquiring points for certification; possible changes in energy use after the official performance period for LEED certification ended; and improvements in the energy efficiency of all federal buildings.

Our findings have important policy implications. The Biden administration plan considers "dramatic investments in energy efficiency in buildings, including completing

4 million retrofits."[29] Attention should be paid to designing policies that provide the intended incentives. In particular, although the USGBC has been revising the LEED scorecard to increase emphasis on the energy component, the trade-offs inherent in LEED certification are likely to remain.[30] If energy efficiency is the primary policy goal, LEED certification may not be the most effective means to reach that goal.

[29] See joebiden.com/clean-energy/.

[30] In 2016, for example, USGBC members voted to increase the minimum simulated improvement in energy performance for new buildings from 10 to 18%, and for major renovations to existing buildings from 5 to 14% (see usgbc.org/credits/core-shell-new-construction/v2009/eap2). Interestingly, they state that "[t]he proposed changes are based on user and volunteer feedback" (see usgbc.org/articles/leed-2009-energy-performance-update). It may well be the case that users were already noticing in practice what we have found in this study.

References

Abadie, Alberto. 2020. "Statistical Nonsignificance in Empirical Economics." *American Economic Review: Insights*, 2(2): 193–208.

Allcott, Hunt, and Michael Greenstone. 2012. "Is There an Energy Efficiency Gap?" *Journal of Economic Perspectives*, 26(1): 3–28.

Blonz, Joshua. forthcoming. "The Costs of Misaligned Incentives: Energy Inefficiency and the Principal-Agent Problem." *American Economic Journal: Economic Policy*.

Brolinson, Becka, Karen Palmer, and Margaret Walls. 2023. "Does Energy Star Certification Reduce Energy Use in Commercial Buildings?" *Journal of the Association of Environmental and Resource Economists*, 10(1): 55–93.

Caliendo, Marco, and Sabine Kopeinig. 2008. "Some Practical Guidance for the Implementation of Propensity Score Matching." *Journal of Economic Surveys*, 22(1): 31–72.

Christensen, Peter, Paul Francisco, Erica Myers, and Mateus Souza. 2023. "Decomposing the Wedge Between Projected and Realized Returns in Energy Efficiency Programs." *Review of Economics and Statistics*.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review*, 110(9): 2964–96.

Eichholtz, Piet, Nils Kok, and John M. Quigley. 2010. "Doing Well by Doing Good? Green Office Buildings." *American Economic Review*, 100(5): 2492–2509.

Eichholtz, Piet, Nils Kok, and John M. Quigley. 2013. "The Economics of Green Buildings." *Review of Economics and Statistics*, 95(1): 50–63.

Fowler, Kimberly M, Emily M Rauch, Jordan W Henderson, and Angela R Kora. 2010. "Re-assessing Green Building Performance: A Post Occupancy Evaluation of 22 GSA buildings." Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram. 2018. "Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program." *Quarterly Journal of Economics*, 133(3): 1597–1644.

F.R., Federal Register. 2007. "Energy Independence and Security Act of 2007." *Public law*.

F.R., Federal Register. 2015. "Executive Order 13693–Planning for Federal Sustainability in the Next Decade." *Federal Register – Presidential Documents*, 80(57): 15871–15884.

Gerarden, Todd D., Richard G. Newell, and Robert N. Stavins. 2017. "Assessing the Energy-Efficiency Gap." *Journal of Economic Literature*, 55(4): 1486–1525.

Gillingham, Kenneth, Amelia Keyes, and Karen Palmer. 2018. "Advances in Evaluating Energy Efficiency Policies and Programs." *Annual Review of Resource Economics*, 10(1): 511–532.

Gillingham, Kenneth, Matthew J Kotchen, David S Rapson, and Gernot Wagner. 2013. "The Rebound Effect is Overplayed." *Nature*, 493(7433): 475–476.

Gillingham, Kenneth, Richard G. Newell, and Karen Palmer. 2006. "Energy Efficiency Policies: A Retrospective Examination." *Annual Review of Environment and Resources*, 31(1): 161–192.

Gillingham, Kenneth, Richard G. Newell, and Karen Palmer. 2009. "Energy Efficiency Economics and Policy." *Annual Review of Resource Economics*, 1(1): 597–620.

Giraudet, Louis-Gaëtan, Sébastien Houde, and Joseph Maher. 2018. "Moral Hazard and the Energy Efficiency Gap: Theory and Evidence." *Journal of the Association of Environmental and Resource Economists*, 5(4): 755–790.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with Variation in Treatment Timing." *Journal of Econometrics*, 225(2): 254–277.

Kahn, Matthew E., Nils Kok, and John M. Quigley. 2014. "Carbon Emissions from the Commercial Building Sector: The Role of Climate, Quality, and Incentives." *Journal of Public Economics*, 113: 1–12.

Kline, Patrick. 2012. "The Impact of Juvenile Curfew Laws on Arrests of Youth and Adults." *American Law and Economics Review*, 14(1): 44–67.

Levinson, Arik. 2016. "How Much Energy Do Building Energy Codes Save? Evidence from California Houses." *American Economic Review*, 106(10): 2867–94.

Lockwood, Charles. 2009. "Building Retro." *Urban Land Magazine*.

Marcus, Michelle, and Pedro H. C. Sant'Anna. 2021. "The Role of Parallel Trends in Event Study Settings: An Application to Environmental Economics." *Journal of the Association of Environmental and Resource Economists*, 8(2): 235–275.

Matisoff, Daniel C, Douglas S Noonan, and Anna M Mazzolini. 2014. "Performance or Marketing Benefits? The Case of LEED Certification." *Environmental Science & Technology*, 48(3): 2001–2007.

Matisoff, Daniel C, Douglas S Noonan, and Mallory E Flowers. 2016. "Policy Monitor—green Buildings: Economics and Policies." *Review of Environmental Economics and Policy*.

McCrary, Justin. 2007. "The Effect of Court-ordered Hiring Quotas on the Composition and Quality of Police." *American Economic Review*, 97(1): 318–353.

Myers, Erica. 2019. "Are Home Buyers Inattentive? Evidence from Capitalization of Energy Costs." *American Economic Journal: Economic Policy*, 11(2): 165–88.

Obama, Barack. 2011. "Implementation of Energy Savings Projects and Performance-Based Contracting for Energy Savings." *Presidential Memorandum.* Available at https://obamawhitehouse.archives.gov/the-press-office/2011/12/02/presidential-memorandum-implementation-energy-savings-projects-and-perfo, accessed on 2018-12-19.

Pombo, Richard W., and William M. Thomas. 2005. "Energy Policy Act of 2005." Available at https://www.govinfo.gov/content/pkg/PLAW-109publ58/html/PLAW-109publ58.htm, accessed on 2019-02-21.

Qiu, Yueming, and Matthew E. Kahn. 2019. "Impact of Voluntary Green Certification on Building Energy Performance." *Energy Economics*, 80: 461–475.

Rosa, Megan. 2016. "Looking Back: LEED History." Available at https://sigearth.com/leed-history/, accessed on 2021-01-10.

Rysman, Marc, Timothy Simcoe, and Yanfei Wang. 2020. "Differentiation Strategies in the Adoption of Environmental Standards: LEED from 2000 to 2014." *Management Science.*

Scofield, John H. 2009. "Do LEED-certified Buildings Save Energy? Not Really. . . ." *Energy and Buildings*, 41(12): 1386–1390.

Scofield, John H. 2013. "Efficacy of LEED-certification in Reducing Energy Consumption and Greenhouse Gas Emission for Large New York City Office Buildings." *Energy and Buildings*, 67: 517–524.

Scofield, John H, and Jillian Doane. 2018. "Energy Performance of LEED-certified Buildings from 2015 Chicago Benchmarking Data." *Energy and Buildings*, 174: 402–413.

Selvacanabady, A, and KS Judd. 2017. *The Influence of Occupancy on Building Energy Use Intensity and the Utility of an Occupancy-Adjusted Performance Metric.* Pacific Northwest National Laboratory.

Sharp, Philip. 1992. "Energy Policy Act of 1992." Available at https://afdc.energy.gov/files/pdfs/2527.pdf, accessed on 2019-02-21.

U.S. DOE. 2020. "About the Commercial Buildings Integration Program." Available at https://www.energy.gov/eere/buildings/about-commercial-buildings-integration-program, accessed on 2019-04-30.

U.S. EIA. 2016. "Recent Energy Intensity Decline in Government Buildings Exceeds Commercial Sector Average." Available at https://www.eia.gov/todayinenergy/detail.php?id=27972, accessed on 2019-04-30.

U.S. EIA. 2020. "Electric Power Monthly with Data for October 2020." Available at https://www.eia.gov/electricity/monthly/, accessed on 2021-01-10.

U.S. EIA. 2021. "How Much Energy is Consumed in U.S. Buildings?" Available at https://www.eia.gov/tools/faqs/faq.php?id=86&t=1, accessed on 2021-02-24.

U.S. EIA. 2022. "2018 Commercial Buildings Energy Consumption Survey." Available at `https://www.eia.gov/consumption/commercial/`, accessed on 2022-08-01.

U.S. EPA. 2019. "Tecnical Reference." Available at `https://portfoliomanager.energystar.gov/pdf/reference/Source%20Energy.pdf`, accessed on 2020-02-02.

U.S. EPA. 2020. "The Difference between Source and Site Energy." Available at `https://www.energystar.gov/buildings/facility-owners-and-managers/existing-buildings/use-portfolio-manager/understand-metrics/difference`, accessed on 2020-01-27.

USGBC, U.S. Green Building Council. 2018. "LEED v4 for Building Operations and Maintenance." Available at `https://www.usgbc.org/resources/leed-v4-building-operations-and-maintenance-current-version`, accessed on 2020-02-01.

USGBC, U.S. Green Building Council. 2020*a*. "LEED Rating System." Available at `https://www.usgbc.org/leed`, accessed on 2020-02-19.

USGBC, U.S. Green Building Council. 2020*b*. "Why LEED." Available at `https://www.usgbc.org/leed/why-leed`, accessed on 2021-01-10.

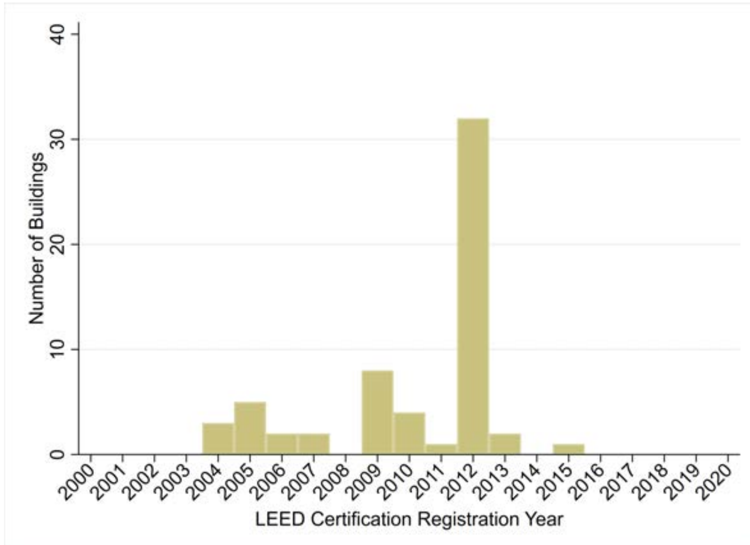
U.S. GSA, U.S. General Services Administration. 2013. "Letter to the U.S. Secretary of Energy." Available at `https://www.gsa.gov/cdnstatic/GSA_Green_Building_Certification_Systems_Review_Letter_to_Sec_Energy.pdf`, accessed on 2019-02-21.

Wedding, G Christopher, and Douglas Crawford-Brown. 2008. "Improving the Link between the LEED Green Building Label and a Building's Energy-related Environmental Metrics." *Journal of Green Building*, 3(2): 85–105.

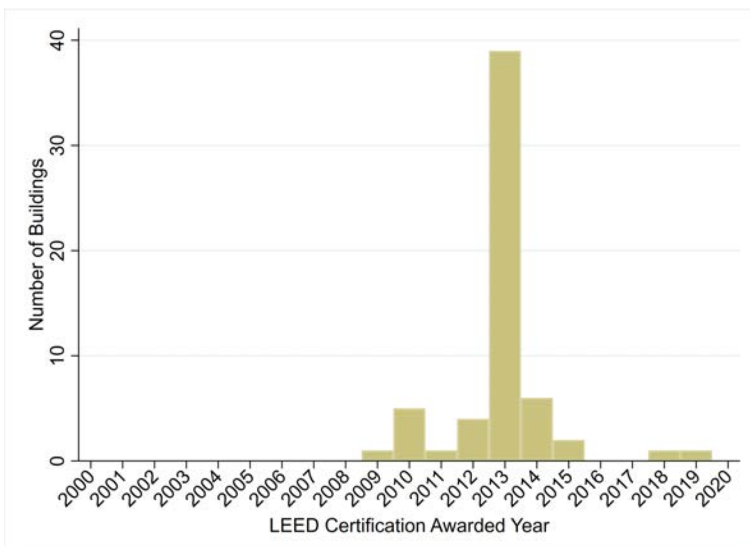
Wehe, Stacey. 2020. "Comparison of LEED Rating Systems." Available at `https://www.redvector.com/lms/userfiles/files/LEED_rating_systems.pdf`, accessed on 2020-02-18.

Figures and Tables

Figure 1: LEED Certification Registration and Awarded Year



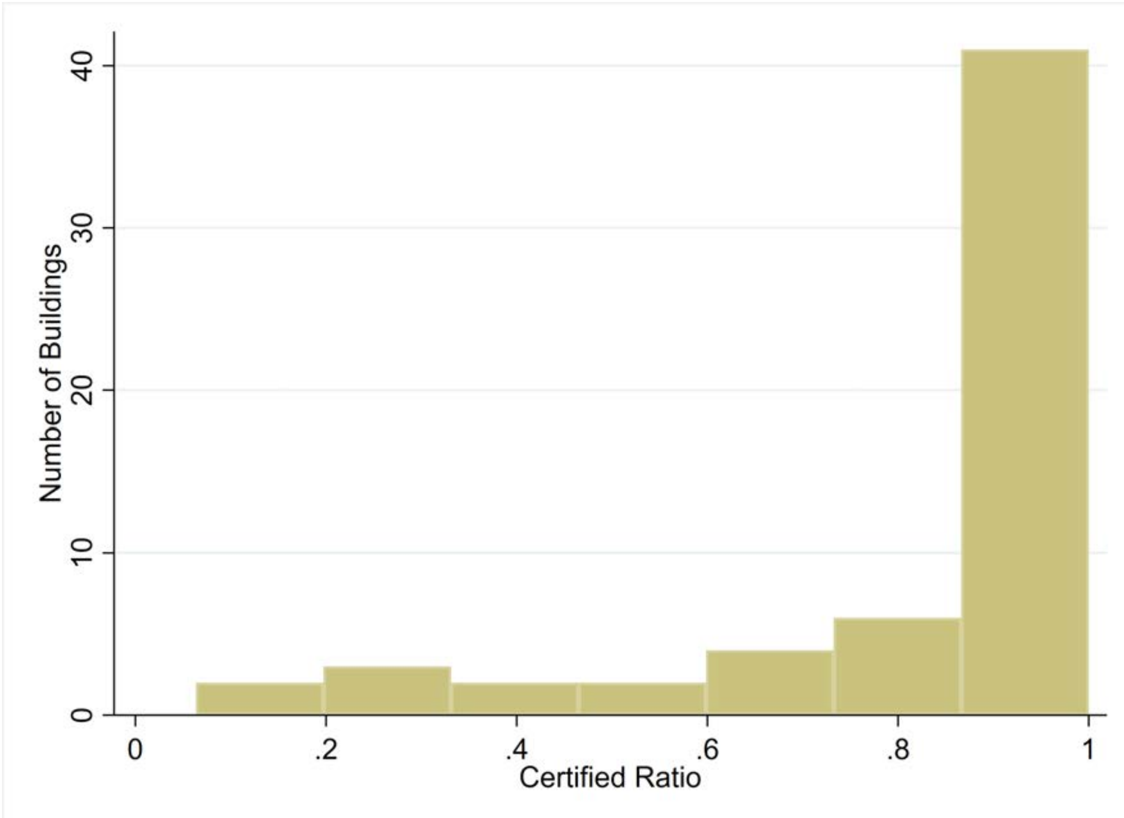
(a) LEED Registration Year



(b) LEED Awarded Year

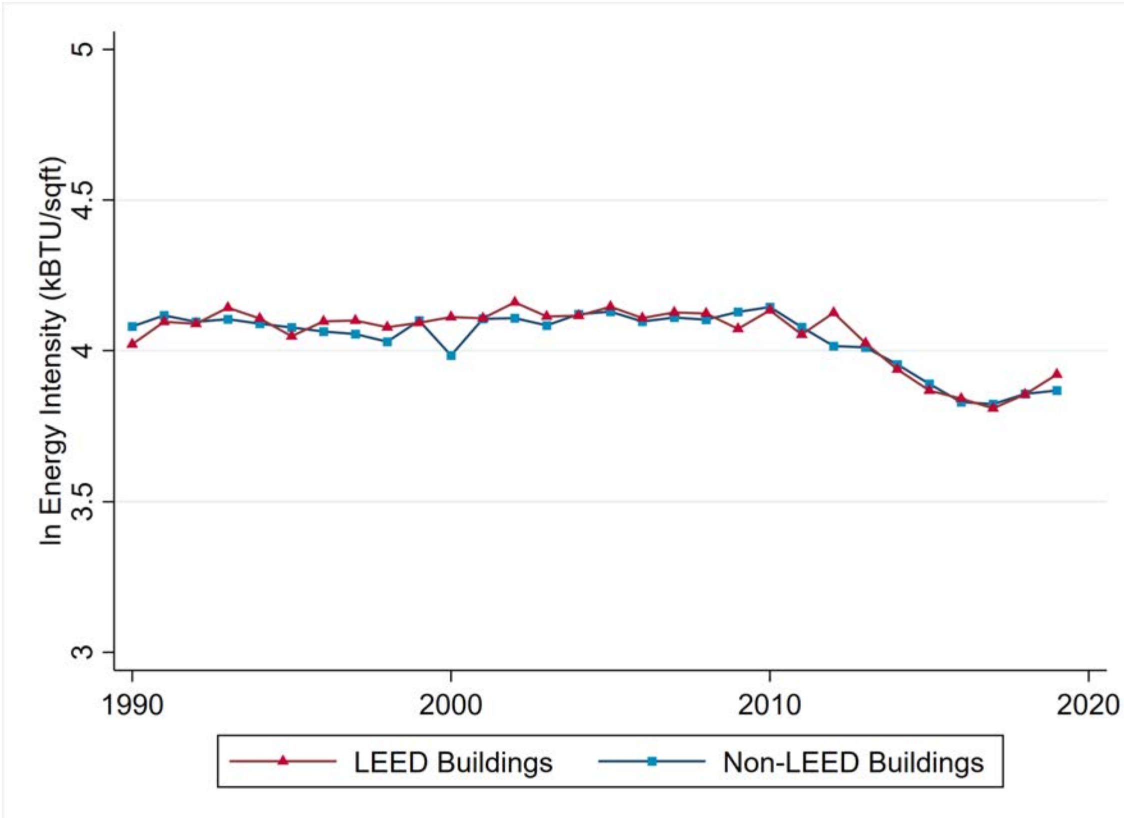
Notes: This figure shows the registered year and the awarded year of the 60 LEED-certified retrofitted buildings in our sample.

Figure 2: The Distribution of the Certified Ratio in the Treated Group



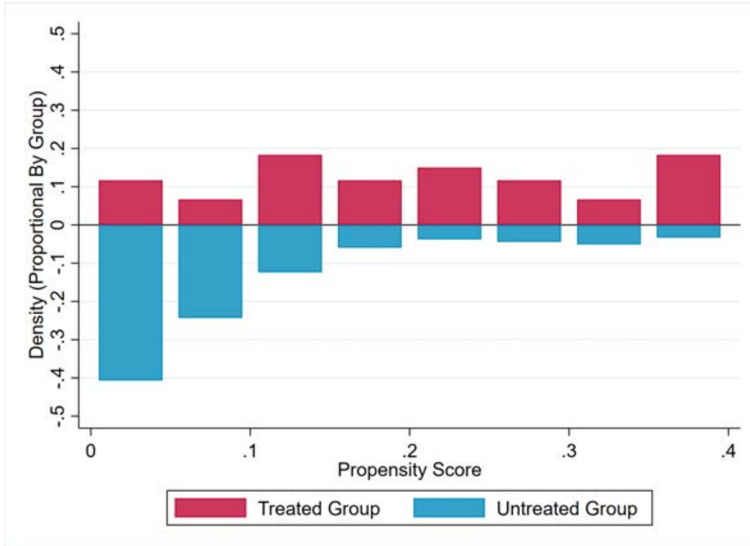
Notes: This figure displays the distribution of the fraction of a building square footage that is LEED-certified – the "certified ratio" – among the LEED buildings in our sample. Note that we regard a building as non-certified unless the ratio is greater than 5%. Thus, the start point of the first bin is 0.05. It is also important to note that existing buildings can obtain LEED certification a new construction if major renovation exceeding 60% of the entire square footage. The ratio average, 10th, 25th, 50th, 75th, and 90th percentile are 0.833, 0.411, 0.799, 0.935, 0.991, and 1.

Figure 3: Energy Use Intensity in Logarithmic Scale of LEED and Non-LEED Buildings

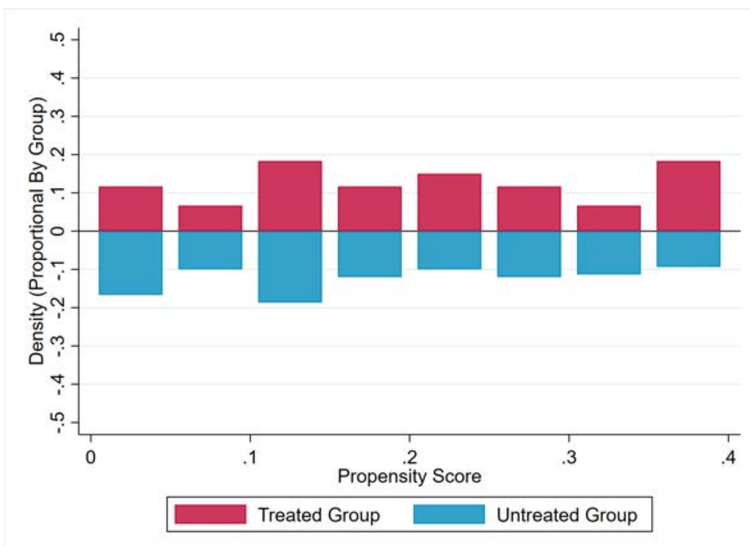


Notes: This figure shows the energy use intensity changes of LEED and non-LEED over the years. The energy consumption is consistent across both groups, and a similar decline in the consumption began around 2010.

Figure 4: Propensity Score Distributions Between Treated and Untreated Buildings



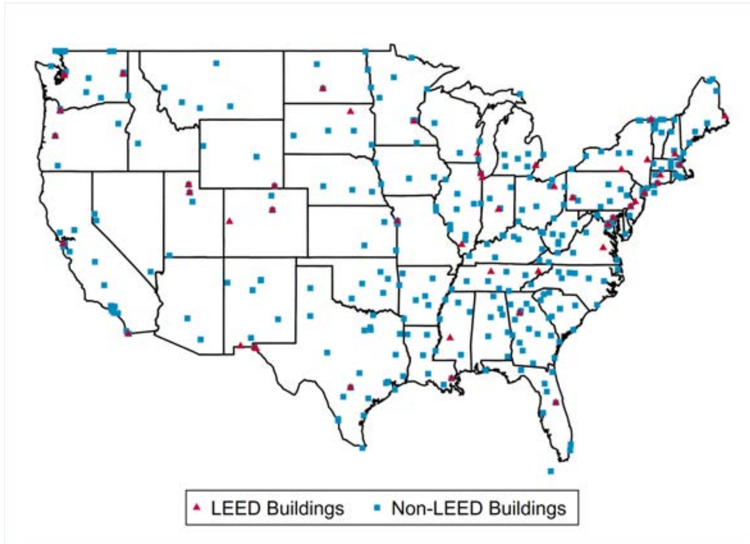
(a) *Before* Propensity Score Matching



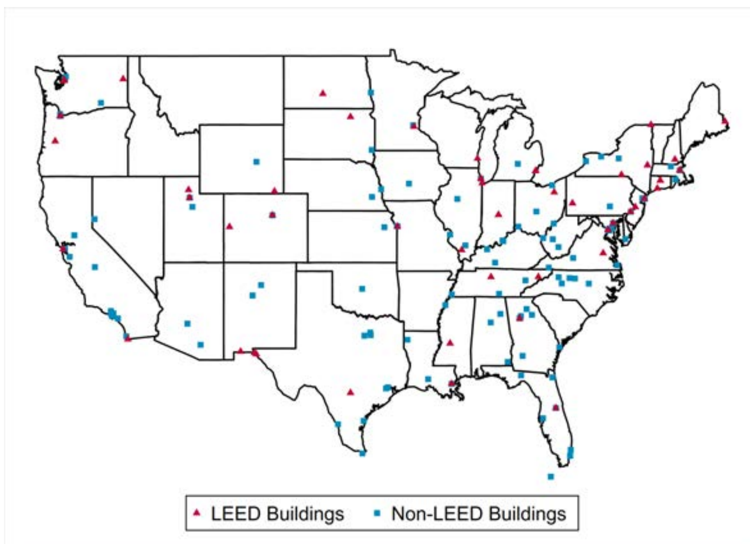
(b) *After* Propensity Score Matching

Notes: This figure presents the distributions of the propensity score for the LEED (red bars) and non-LEED (blue bars) federal buildings before and after matching. The horizontal axis displays the propensity score, and the vertical axis the proportion of buildings that falls into each propensity score bin. (a) shows the score distribution before matching; (b) presents the distribution after five-nearest-neighbor (5NN) propensity score matching with replacement. As observed, the propensity score distributions are similar between the two groups after matching in (b). The symmetric pattern between groups suggests that the common support assumption might hold in our analysis. Furthermore, our matching results also indicate that each individual LEED building is able to find a match from the untreated group.

Figure 5: Locations of LEED and Non-LEED Buildings



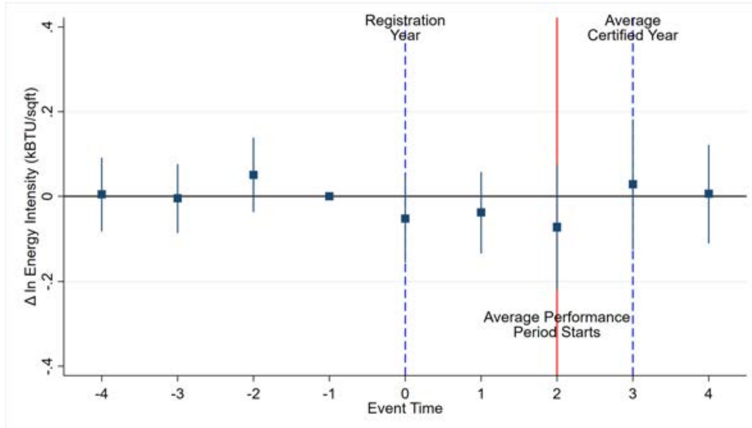
(a) Before Propensity Score Matching



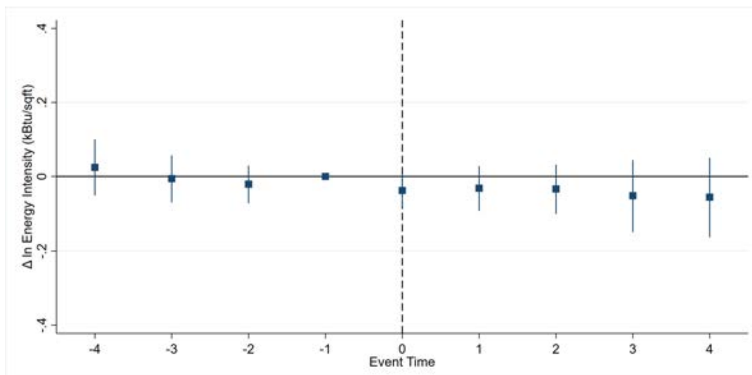
(b) After Propensity Score Matching

Notes: This figure maps the locations of the LEED and non-LEED buildings before and after the propensity score matching. (a) maps the location of the buildings before matching; (b) maps the building locations after matching with 5NN.

Figure 6: Energy Savings in Each Period Relative to the Savings in the Reference Year



(a) Event Study Estimates



(b) Falsification Test on Non-LEED Buildings

Notes: This figure displays estimates from event study analyses for LEED certification among federal buildings. The vertical axis of the graphs displays the log difference between energy use intensity (in kBTU/ft^2) of treated and untreated federal buildings. (a) shows the event study results by plotting the coefficients δ's from Equation (3). The reference year is set to be one year prior to the registration date. Buildings are on average certified three years after they registered with LEED, and a building operation monitor period, "performance period", starts one year before certification is eventually awarded. (b) provides the falsification test results with the non-LEED federal buildings only. For this test, we assign a treatment indicator to the untreated buildings based on their sizes, when compared to their LEED-certified counterparts. Detailed procedures of the test are discussed in Appendix III. The vertical bar around each estimated coefficient indicates the 95% confidence interval.

Table 1: LEED Component Credit Category

Score Components	Major Credit Categories	Points	Examples
Energy and Atmosphere	Optimize Energy Performance	18	High efficiency lighting with daylighting controls, high efficiency HVAC, high performance glazing and slab insulation
	Renewable Energy	6	Green power
	Building Commissioning	4	Energy audit
Sustainable Sites	Alternative Commuting Options	15	Public bus lines, campus shuttle buses
Water Efficiency	Indoor Fixture Efficiency	5	Water saving sink faucets, WaterSense certified automatic flush
	Water Efficient Landscaping	5	High efficiency irrigation technology, Sprinkler pump
Material and Resources	Sustainable Purchasing	6	Sustainable furniture, reduced mercury in lamps, bathroom hand dryers
	Solid Waste Management	4	Recycled materials, waste stream audit
Indoor Environmental Quality	Green Cleaning	6	Sustainable cleaning products and materials
	Indoor Air Quality Best Management Practices	5	Effective ventilation, filter installation
Innovation in Operations	Innovation in Operations	4	Innovations on top of LEED requirements
	LEED Accredited Professional	1	Hire a LEED AP

Notes: This table provides an overview of some important credit categories for each component in the LEED scorecard using LEED EB 2009 as an example. LEED EB means LEED for existing buildings. More than half of the treated buildings included in our sample were certified based on LEED EB 2009, as reported in Appendix Table A.2. Note that the credit categories listed here are only the ones that contain the majority of points. The total achievable score for LEED EB 2009 is 110, with 35 points for Energy and Atmosphere, 26 points for Sustainable Sites, 14 points for Water Efficiency, 10 points for Material and Resources, 15 points for Indoor Environmental Quality, 6 points for Innovation in Operations, and 4 points for Regional Priority. Our study does not include the discussion of Regional Priority because it grants points based on local issues, thus its requirements vary across different LEED programs. GBIG reports do not record the points for this component either.

Table 2: Descriptive Statistics for Buildings With and Without LEED Certification

Panel A: Building Characteristics	Pre-treatment Site log(EUI) (kBTU/ft^2)	Pre-treatment Source log(EUI) (kBTU/ft^2)	Building Size (ft^2)
Treated Buildings (Average)	4.093	4.863	409,851
	[0.547]	[0.513]	[304,934]
Untreated Before Matching	4.077	4.830	217,808
	[0.744]	[0.699]	[349,405]
Difference Before Matching	0.016	0.033	192,043***
	(0.079)	(0.074)	(42,660)
Untreated After 5NN Matching	4.059	4.904	428,236
	[0.597]	[0.559]	[306,521]
Difference After 5NN Matching	0.035	-0.041	18,385
	(0.079)	(0.076)	(43,162)

Panel B: Number of Buildings	Office	Courthouse	Other	Total
Treated Buildings	45	12	3	60
Untreated Before Matching	295	122	35	452
Untreated After 5NN Matching	211	83	6	300

Panel C: LEED Attributes	CR	ES	SS	WS	MIIS
Treated Buildings (Average)	0.833	0.155	0.112	0.050	0.193
	[0.239]	[0.069]	[0.060]	[0.027]	[0.085]

Notes: This table reports average characteristics of federal buildings with LEED certification and those without it, *before* and *after* five-nearest-neighbor (5NN) matching. The sample includes buildings observed over the period 1990-2019. Panel A reports building characteristic variables for buildings in the treated group and buildings in the untreated group before and after 5NN matching. Pre-treatment is defined as the first 12 months that a building appears in the dataset. EUI stands for energy use intensity, which is defined as the ratio of total energy consumption and the square footage of a building. Site energy is the amount of heat and electricity consumed by a building as reflected in the utility bill, whereas source energy represents the total amount of raw fuel that is required to operate a building, incorporating all transmission, delivery, and production losses. Panel B shows the number of buildings for treated and untreated buildings before and after matching. It is important to note that the untreated buildings are matched to LEED buildings with replacement, thus non-LEED buildings can appear more than once in the untreated group after 5NN matching. Panel C presents averages of different attributes in the LEED program for LEED-certified buildings. CR stands for certified ratio, and refers to the fraction of a building square footage that is LEED-certified. ES, SS, WS, and MIIS represent the component score proportions evaluated by the LEED program: energy and atmosphere, sustainable sites, water efficiency, and the combination of material, indoor environment, and innovation, respectively. Score proportion is defined as the ratio of the component score and the total achievable score in the LEED program. Standard deviations are reported in brackets, and standard errors clustered by building in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 3: Descriptive Regression of Energy Use Intensity on LEED Program Attributes

	(1) All Buildings	(2) Office
CR	-0.040	-0.005
	(0.066)	(0.070)
ES	-0.308***	-0.309***
	(0.072)	(0.084)
WS	0.134**	0.173**
	(0.059)	(0.074)
SS	-0.011	0.023
	(0.059)	(0.060)
MIIS	-0.174**	-0.183**
	(0.086)	(0.090)
R-squared	0.326	0.387
Observations	60	45

Notes: This table reports OLS regression results of energy use intensity on different attributes in the LEED program, including certified ratio and LEED component scores during the year when LEED buildings were certified. The dependent variable is the log of site energy use intensity in the LEED certified years. CR stands for certified ratio, and refers to the fraction of a building square footage that is LEED-certified. ES, SS, WS, and MIIS represent the component scores evaluated by the LEED program: energy and atmosphere, sustainable sites, water efficiency, and the combination of material, indoor environment, and innovation, respectively. To facilitate comparison, CR, ES, SS, WS, and MIIS are all standardized to have mean zero and standard deviation one. Column (1) includes all federal buildings, and column (2) includes office buildings only. As observed, energy retrofit is positively correlated with energy use intensity while water retrofit is negatively correlated with energy use intensity. Standard errors clustered by building are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 4: Energy Use Intensity Impacts from the LEED Program and Heterogeneity Analysis Regarding Different Program Attributes

	(1) All Buildings	(2) All Buildings	(3) Office	(4) All Buildings	(5) Office	(6) All Buildings	(7) Office
T	-0.030	-0.038	0.001	-0.035	0.010	-0.039	0.007
	(0.045)	(0.040)	(0.051)	(0.040)	(0.051)	(0.040)	(0.050)
T × CR		-0.054*	-0.041	-0.055**	-0.041	-0.060**	-0.047
		(0.028)	(0.030)	(0.028)	(0.030)	(0.027)	(0.030)
T × ES		-0.126***	-0.139***	-0.122***	-0.133***	-0.122***	-0.135***
		(0.037)	(0.048)	(0.035)	(0.046)	(0.035)	(0.046)
T × SS		0.024	0.038	0.022	0.034	0.022	0.036
		(0.046)	(0.051)	(0.045)	(0.050)	(0.045)	(0.050)
T × WS		0.063*	0.083**	0.059*	0.073**	0.060*	0.072**
		(0.033)	(0.035)	(0.033)	(0.034)	(0.033)	(0.034)
T × MIIS		-0.008	-0.015	-0.012	-0.019	-0.015	-0.025
		(0.032)	(0.037)	(0.032)	(0.036)	(0.031)	(0.036)
Energy Star				-0.079**	-0.106**	-0.075*	-0.102**
				(0.039)	(0.049)	(0.039)	(0.048)
ECM						-0.063*	-0.067
						(0.035)	(0.042)
Year FE	×	×	×	×	×	×	×
Building FE	×	×	×	×	×	×	×
Adj. R-squared	0.665	0.669	0.565	0.670	0.567	0.670	0.568
Observations	9,599	9,599	6,894	9,599	6,894	9,599	6,894

Notes: This table reports estimates of energy savings from the LEED program in federal buildings over the period 1990-2019, and the results of the heterogeneity analysis regarding different program attributes with and without the control of other energy conservation measures. The dependent variable is the log of site energy use intensity, and the untreated group is constructed using five-nearest-neighbor (5NN) matching with replacement. Energy use intensity is defined as the ratio of total energy consumption and the square footage of a building. Site energy is the amount of heat and electricity consumed by a building as reflected in the utility bill. T is the treatment indicator equal to one if a building i in year t is registered with LEED, and zero otherwise. CR stands for certified ratio, and refers to the fraction of a building square footage that is LEED-certified. ES, SS, WS, and MIIS represent the component scores evaluated by the LEED program: energy and atmosphere, sustainable sites, water efficiency, and the combination of material, indoor environment, and innovation, respectively. To facilitate comparison, CR, ES, SS, WS, and MIIS are all standardized to have mean zero and standard deviation one. Energy Star is the treatment indicator for Energy Star program equal to one of a building i was labeled in year t. ECM is the treatment indicator for efficiency conservation measures equal to one if a building i completed the conservation program in year t. Column (1) reports the aggregate energy savings from the LEED program. Columns (2) and (3) show estimates when including all program attributes in the regression. Columns (4) and (5) add controls for the Energy Star program. Columns (6) and (7) estimate the specification with the inclusion of controls for both Energy Star and other efficiency conservation upgrades. Columns (1), (2), (4), and (6) report the results for all buildings, and columns (3), (5), and (7) restrict attention to office buildings. The coefficients on LEED certification in columns (1)-(7) are consistently small and insignificant, which suggest that LEED certification did not have a statistically significant effect on energy consumption. A significant negative estimate on energy component score, and a significant positive estimate on water component score after controlling for the rest of the attributes in the LEED program in columns (2)-(7) suggest a trade-off between energy and water components when a building is LEED certified. CDD and HDD are included as controls in all regressions. Because of the potentially repeated non-LEED buildings in the untreated group due to matching with replacement, all regressions are weighted by the inverse of the number of repeated observations. Standard errors clustered by building are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 5: Energy Use Intensity Impacts from the LEED Program with Different Specification and Sample Restrictions

	D.V.: EUI Levels		Certification Awarded as Treatment		W/O Matching		Matching Within Census Regions		Matching Within Census Divisions	
	(1) All Buildings	(2) Office	(3) All Buildings	(4) Office	(5) All Buildings	(6) Office	(7) All Buildings	(8) Office	(9) All Buildings	(10) Office
T	1.547	4.592	-0.023	0.007	-0.044	-0.005	-0.028	0.011	-0.002	0.027
	(2.978)	(3.489)	(0.049)	(0.059)	(0.039)	(0.048)	(0.037)	(0.046)	(0.040)	(0.050)
T × CR	-4.342*	-3.418	-0.014	0.009	-0.054*	-0.041	-0.067**	-0.053*	-0.101***	-0.084**
	(2.328)	(2.471)	(0.043)	(0.045)	(0.028)	(0.030)	(0.028)	(0.030)	(0.034)	(0.039)
T × ES	-8.261**	-8.175*	-0.163***	-0.176***	-0.125***	-0.138***	-0.114***	-0.119**	-0.098***	-0.092**
	(3.490)	(4.261)	(0.048)	(0.059)	(0.036)	(0.047)	(0.036)	(0.046)	(0.031)	(0.045)
T × SS	2.643	3.223	0.037	0.061	0.024	0.036	-0.014	-0.004	-0.078*	-0.088*
	(2.389)	(2.577)	(0.049)	(0.053)	(0.046)	(0.051)	(0.034)	(0.039)	(0.043)	(0.051)
T × WS	6.102**	7.739**	0.082**	0.131***	0.063*	0.082**	0.066*	0.093**	0.110***	0.142***
	(2.779)	(3.149)	(0.036)	(0.039)	(0.033)	(0.034)	(0.034)	(0.037)	(0.037)	(0.043)
T × MIIS	0.213	0.642	0.011	0.002	-0.010	-0.016	-0.004	-0.006	0.038	0.054
	(2.567)	(2.711)	(0.046)	(0.048)	(0.031)	(0.036)	(0.030)	(0.035)	(0.037)	(0.044)
Year FE	×	×	×	×	×	×	×	×	×	×
Building FE	×	×	×	×	×	×	×	×	×	×
Adj. R-squared	0.758	0.695	0.670	0.567	0.798	0.766	0.666	0.690	0.700	0.723
Observations	9,599	6,894	9,599	6,894	13,703	9,255	8,991	6,102	7,250	5,317

Notes: This table reports results from several robustness checks by re-estimating the preferred specifications of columns (2) and (3) of Table 4 with different restrictions. T is the treatment indicator. In columns (1)-(2) and (5)-(10), T equal to one if a building i in year t is registered with LEED, and zero otherwise. In columns (3) and (4), T equal to one if a building i in year t is awarded with LEED certification. CR stands for certified ratio, and refers to the fraction of a building square footage that is LEED-certified. ES, SS, WS, and MIIS represent the component scores evaluated by the LEED program: energy and atmosphere, sustainable sites, water efficiency, and the combination of material, indoor environment, and innovation, respectively. To facilitate comparison, CR, ES, SS, WS, and MIIS are all standardized to have mean zero and standard deviation one. The dependent variable is the site energy use intensity *without* log transformation for columns (1) and (2); the log of site energy use intensity for columns (3)-(10). The untreated group is constructed using five-nearest-neighbor (5NN) matching with replacement for all columns except for columns (5) and (6). Columns (1) and (2) use the site energy use intensity levels as the dependent variable. Columns (3) and (4) use LEED certification awarded year as treatment year instead of LEED program registration year. Columns (5) and (6) include all buildings in the sample and estimate the coefficients without the first stage propensity score. Columns (7) and (8) match LEED buildings with non-LEED buildings within the same Census Regions (4 Census Regions). Columns (9) and (10) match LEED buildings with non-LEED buildings within the same Census Divisions (9 Census Divisions). Columns (1), (3), (5), (7), and (9) include all federal buildings, and columns (2), (4), (6), (8), and (10) include office buildings only. Comparing to the main results in Table 4, the estimates for different attributes of the LEED program are similar across all cases. CDD and HDD are included as controls in all regressions. Because of the potentially repeated non-LEED buildings in the untreated group due to matching with replacement, all regressions except for columns (5) and (6) are weighted by the inverse of the number of repeated observations. Standard errors clustered by building are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 6: Results of the Goodman-Bacon Decomposition for LEED Registration Year

	(1) 1990–2019	(2) 1995–2019	(3) 2000–2019
Overall DD Estimate	-0.018	-0.022	-0.054
	(0.050)	(0.044)	(0.038)
DD Est.: T vs. Never Treated	-0.010	-0.012	-0.043
DD Est.: Earlier T vs. Later C	-0.039	-0.046	-0.068
DD Est.: Later T vs. Earlier C	-0.223	-0.239	-0.253
Weights: T vs. Never Treated	0.917	0.912	0.915
Weights: Earlier T vs. Later C	0.055	0.050	0.036
Weights: Later T vs. Earlier C	0.028	0.037	0.049
Number of LEED Buildings	49	51	53
Number of Non-LEED Buildings	94	101	121
Number of Observations	4,290	3,800	3,480

Notes: This table reports energy savings of the LEED program in federal buildings from running the Goodman-Bacon decomposition. The regressions include building fixed effects and year fixed effects. The Goodman-Bacon method decomposes the overall difference-in-differences ("DD") estimate into three components: (i) buildings that were ever registered versus buildings that were never registered ("T vs. Never Treated"), (ii) buildings that were registered earlier, using buildings registered later as controls ("Earlier T vs. Later C"), and (iii) buildings that were registered later, using buildings registered earlier as controls ("Later T vs. Earlier C"). For each component, the decomposition provides both the DD estimate and the corresponding weight to derive the overall DD estimate. Standard errors clustered by building are reported in parentheses for the overall DD estimate. The decomposition requires a strongly balanced panel and does not allow any weighting options. To construct this, we remove the repeated buildings in the untreated group and only include LEED and non-LEED buildings with *consecutive* observations between time periods specified in column (1), (2), and (3). The overall DD estimate is consistent with the coefficient in the main table and shows zero savings from the LEED program. More importantly, around 93 percent of the variation used to calculate the overall DD estimate is from the comparison between treated and never treated buildings, with no negative weights. This is not surprising because over two thirds of the LEED-certified retrofitted buildings in our sample were registered in one single year, as depicted in Figure 1. Thus, our estimate is unlikely to be biased because of the variation in treatment timing.

Online Appendix (Not For Publication)

"Does LEED Certification Save Energy? Evidence from Federal Buildings"

Appendix I: LEED Program Introduction

LEED was first developed and launched by USGBC with the help from the Federal Energy Management Program in 1998 (Rosa, 2016). Unlike other widely recognized labels such as Energy Star, which focus energy use, LEED certifies a building based on a comprehensive evaluation of six categories. When the total score of these categories exceeds a certain range, the building is LEED certified.

There are several LEED programs included in our dataset (Appendix Table A.2), but more than half of the buildings in our sample were certified with LEED EB 2009 (LEED for Existing Buildings). Note that buildings do not have to be newly constructed to be certified with LEED New Construction, as long as at least 60% of the project's square footage is completed by the certification date. On the other hand, existing buildings can also apply for operations and maintenance if they are fully operational for at least one year.

The LEED scorecard generally includes energy and atmosphere, materials and resources, indoor environmental quality, sustainable sites, water efficiency, and innovation in operations. With each of them taking up different proportions of the total score, the energy component accounts for the largest share, more than 30%.[1] Table 1 sketches some main credit categories for different components with examples. Some upgrades in the non-energy components might result in higher energy consumption. For example, water saving faucets with sensors might consume more energy compared to those without.

[1] There could be minimum score requirement for some components, but the mandatory score is small in comparison to the maximum score of the components.

Appendix II: Data Sources Description

The data used for the analysis in this paper are from five sources. They are publicly available and can be accessed online.

1. EUAS Dataset

   This dataset provides monthly energy consumption data for the General Service Administration (GSA) buildings in the United States.

   `https://catalog.data.gov/dataset/energy-usage-analysis-system`

2. GBIG Reports

   The GBIG website provides detailed information of the LEED certification status of the GSA buildings, including program registration date, certification awarded date, points earned in each component of the LEED scorecard, etc.

   `http://www.gbig.org/collections/14796/buildings?page=3`

3. NOAA Weather Data

   `https://www.cpc.ncep.noaa.gov/products/analysis_monitoring/cdus/degree_days/`

4. GSA Building Address

   This dataset provides building address for each GSA building. The address is used to match LEED information provided in the GBIG reports with energy consumption data in the EUAS dataset.

   `https://www.iolp.gsa.gov/iolp/`

5. GSA Building Efficiency Conservation Measures

   `https://ctsedwweb.ee.doe.gov/CTSDataAnalysis/Reports/PublicReport_ProjectsByFiscalYear.aspx?Covered=0`

6. Registry of Energy Star Certified Buildings and Plants

   `https://ctsedwweb.ee.doe.gov/CTSDataAnalysis/Reports/PublicReport_ProjectsByFiscalYear.aspx?Covered=0`

Appendix III: Falsification Test on Non-LEED Buildings

The energy savings are derived by comparing the difference in consumption between LEED and non-LEED buildings in the post certification period relative to the same difference before the reference year. Although we implemented the propensity score matching, it would still be a concern if the untreated buildings are affected by some shocks that are missed from the difference-in-differences model but would drive down their energy use intensity.

We repeat the first and second stage regressions with only the non-LEED buildings and report the δ's from equation (3). We follow the step below to first assign the treatment indicator to the untreated buildings:

1. Generate bins based on the building size with bandwidth of 7,500 ft^2.
2. If the non-LEED buildings fall into the same bin with a LEED building, those non-LEED buildings are assigned the treatment date of this LEED building.
3. If two or more LEED building appear in the same bin, we generate random numbers for both LEED and non-LEED buildings in the bin. Then, we assign the treatment date of the LEED buildings to the non-LEED buildings if their numbers match.
4. If a non-LEED building lays in a bin that does not contain any LEED buildings, this non-LEED building will not be assigned to any treatment date and is regarded as untreated.

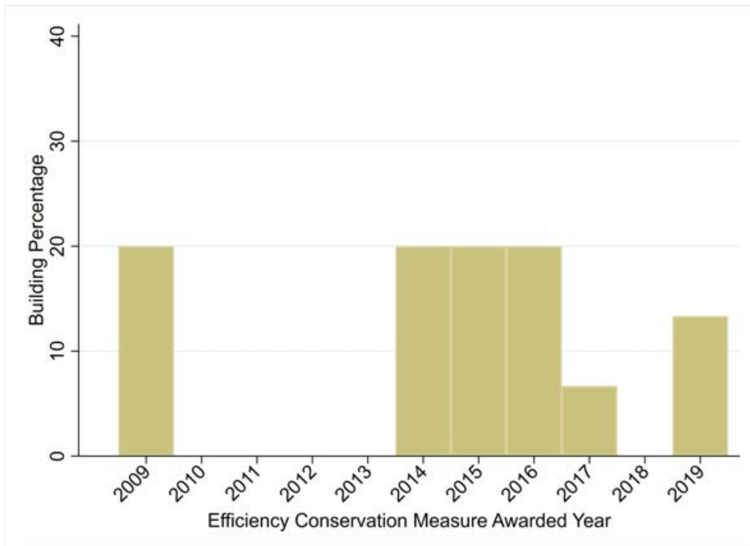
Take the sample after 5NN matching as an example. Among the 300 non-LEED buildings, 142 of them are assigned as treated while all the rest as assigned as untreated. Then, we follow the steps similar to the main results estimation and construct a balanced panel for the treated buildings to ensure enough pre- and post- treatment periods. This process eliminates 2 buildings from the treated group and leaves us with 140 buildings.[2] We then apply the same specification as the main results for the propensity score matching and the event study regression.

[2] Because there are repeated buildings in the non-LEED buildings group from the with replacement propensity score matching stage, the number of buildings identified in the text does not correspond to the amount of unique buildings. These 2 eliminated buildings are from one unique building.

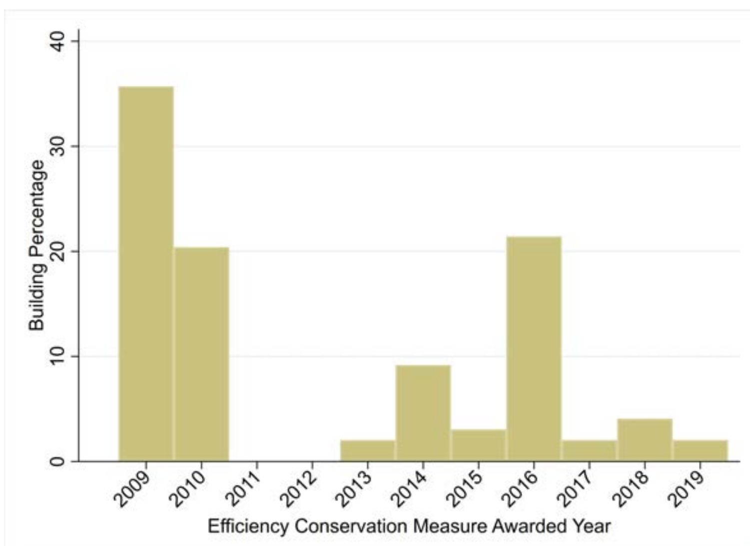
Figure 6b plots the results of the second stage using non-LEED buildings after 5NN matching. All the δ's are around zero in the figure, which suggest that there is no shocks in the non-LEED buildings. We also conduct the robustness checks for the sample after 3NN matching and observe similar zero δ's for all event times (see Appendix Figure A.5).

Appendix IV: Figures and Tables

Figure A.1: Completion Year of the Efficiency Conservation Program for LEED and Non-LEED Buildings



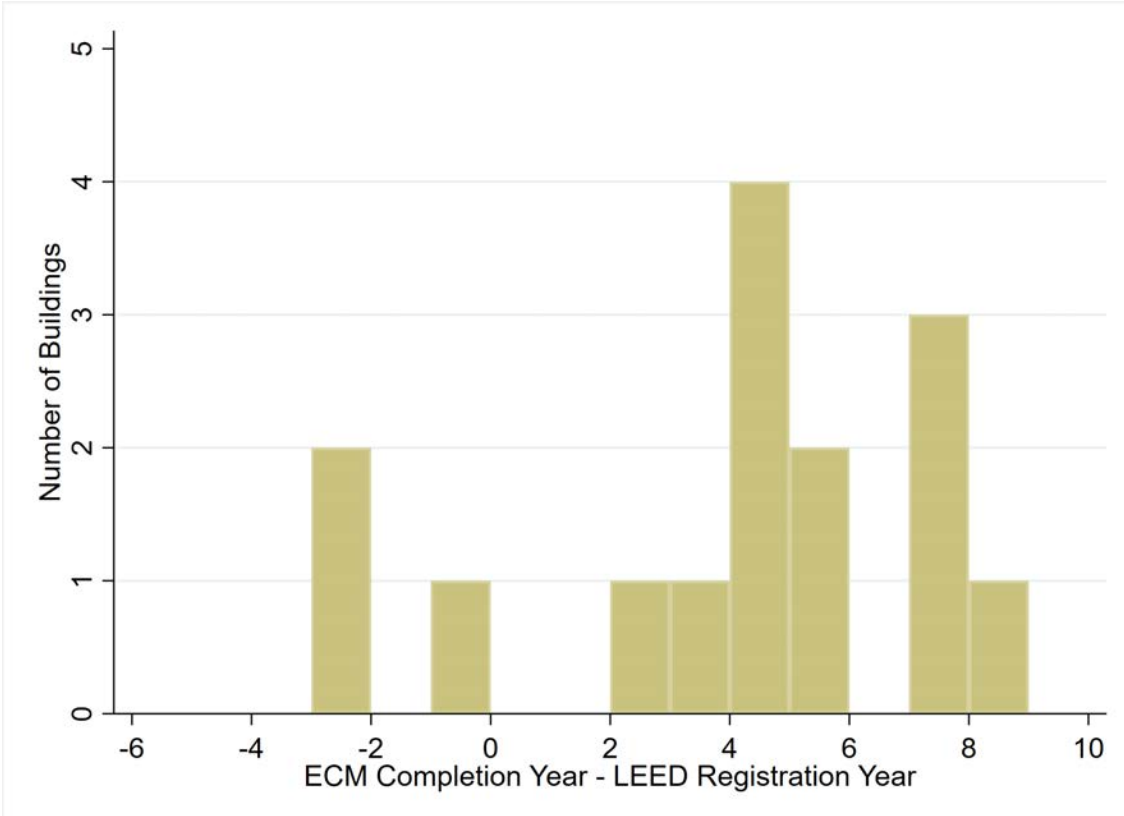
(a) LEED Buildings



(b) Non-LEED Buildings

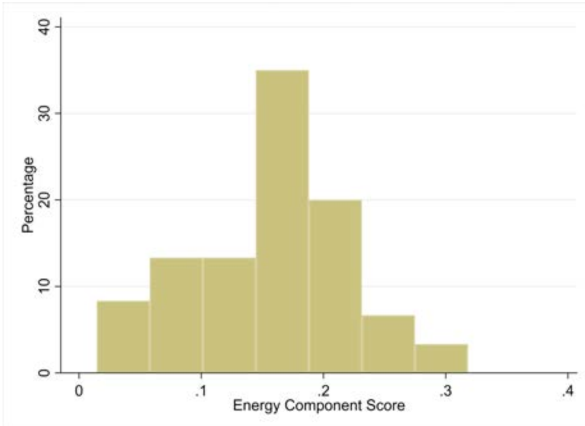
Notes: This figure is the histogram of the completion year of the efficiency conservation program in LEED and non-LEED buildings after 5NN matching.

Figure A.2: The Distribution of the Year Differences of the Efficiency Conservation Upgrade and LEED Certification

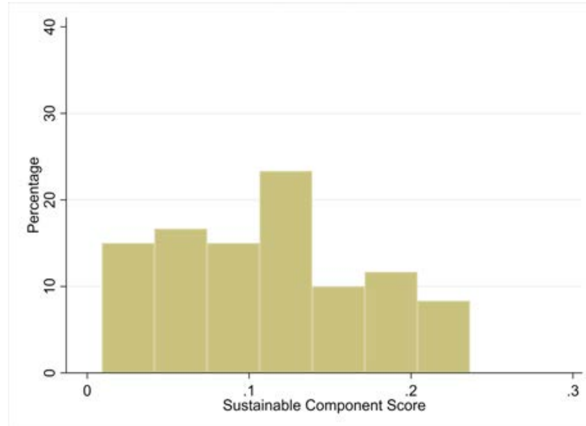


Notes: This figure provides the distribution of the year differences of the efficiency conservation upgrade and the LEED program. For example, 4 means the first energy conservation upgrade in the building was completed 4 years after the building registered for the LEED program. Among the 60 LEED certified buildings, 15 of them have adopted at least one efficiency conservation upgrade. This figure is plotted with these 15 LEED buildings.

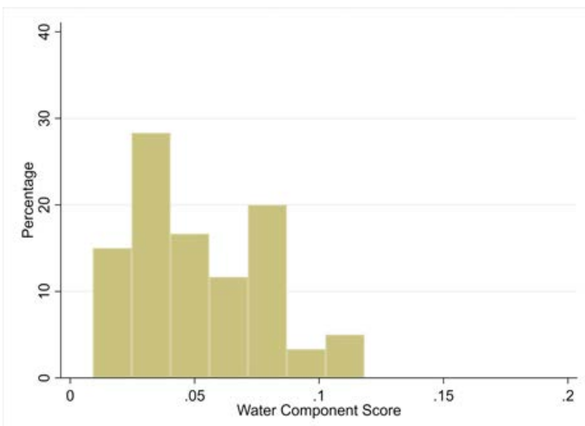
Figure A.3: The Distribution of the Component Score



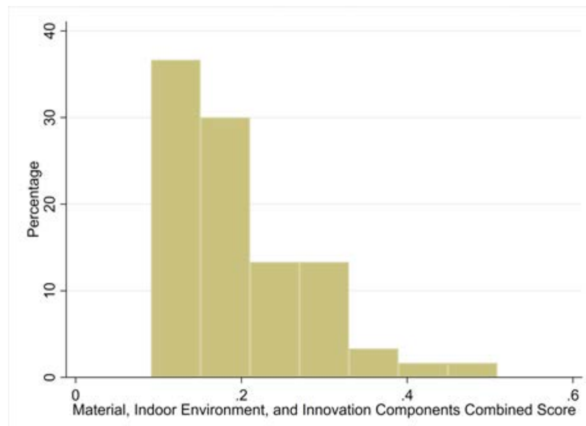
(a) Energy Component Score



(b) Sustainable Component Score



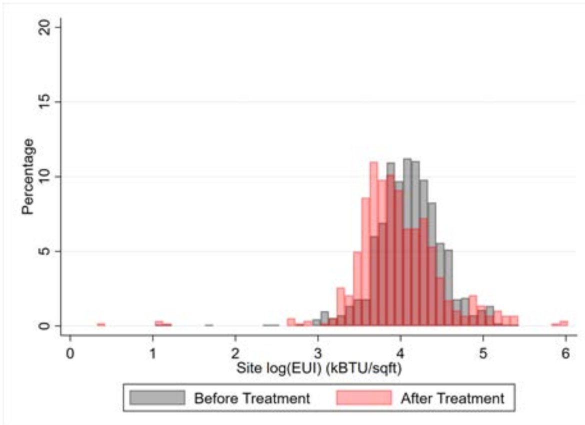
(c) Water Component Score



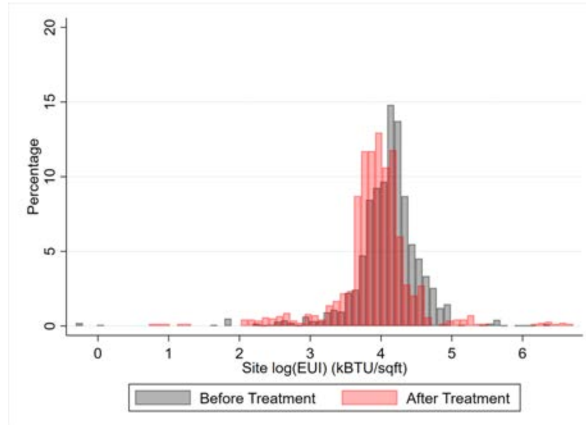
(d) Material, Indoor Environment, and Innovation Components Combined Score

Notes: This figure presents the distributions of the component scores in the LEED buildings. The energy, sustainable and water component scores spread out symmetrically. The material, indoor environment, and innovation components combined score is right skewed with most of the buildings fall in the first bin. The score average, 10th, 50th, and 90th percentile are 0.155, 0.058, 0.159, and 0.232 for the energy component; 0.112, 0.036, 0.112, and 0.200 for the sustainable component; 0.050, 0.016, 0.050, and 0.082 for the water component; 0.193, 0.109, 0.159, and 0.290 for the material, indoor environment, and innovation components combined.

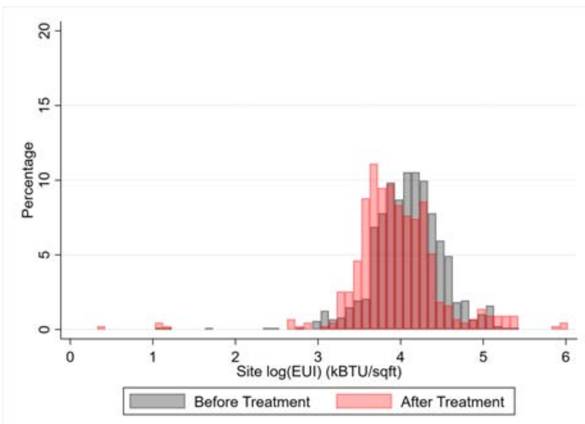
Figure A.4: The Distribution of Average Energy Use Intensity (in Logarithm Form) Before and After the Certification



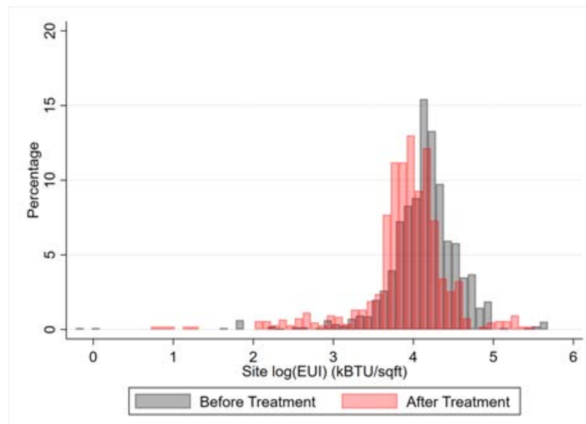
(a) LEED buildings



(b) Non-LEED buildings (Falsification Test)



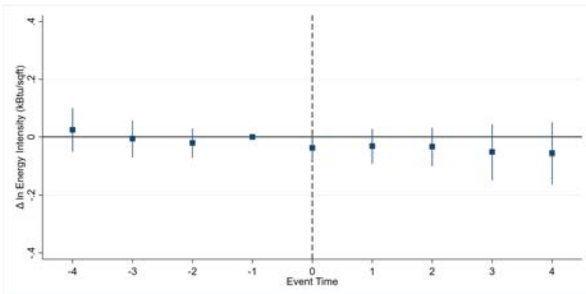
(c) LEED Office buildings



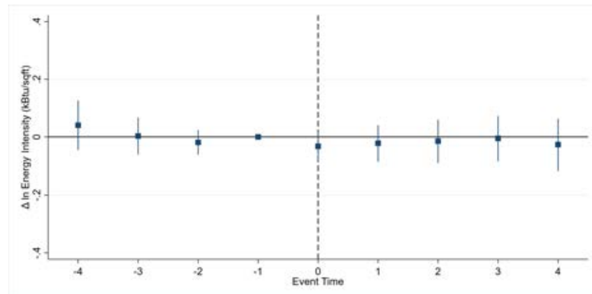
(d) Non-LEED Office buildings (Falsification Test)

Notes: This figure presents the distributions the average energy use intensity in LEED and non-LEED buildings. (a) plots the histogram of average energy use intensity in LEED buildings both before and after the certification. (b) follows the Falsification Test procedures (Appendix III), and depicts the histogram of average energy use intensity in non-LEED buildings before and after the assigned treatment date. Similarly, (c) and (d) are the histograms of LEED and non-LEED office buildings.

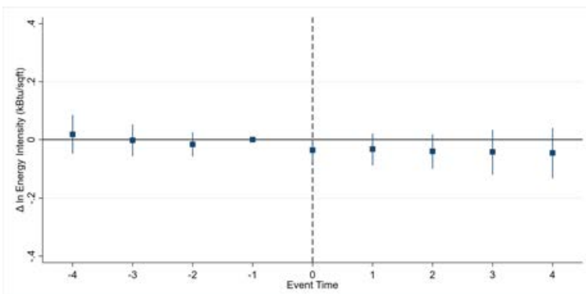
Figure A.5: Energy Savings in Each Period Relative to the Reference Year Savings for the Untreated Buildings Only
(*Falsification Test*)



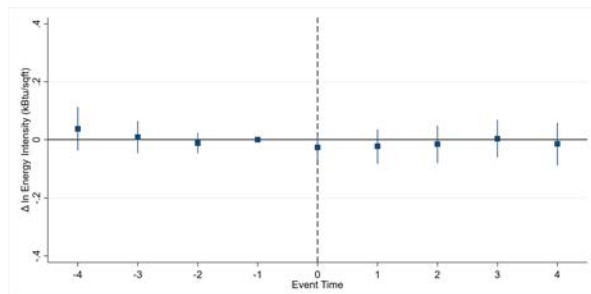
(a) Site EUI with 5NN



(b) Site EUI with 3NN



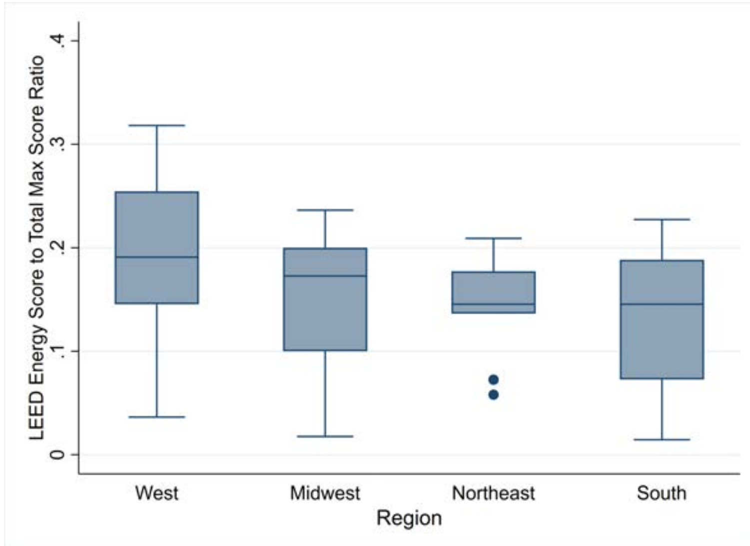
(c) Source EUI with 5NN



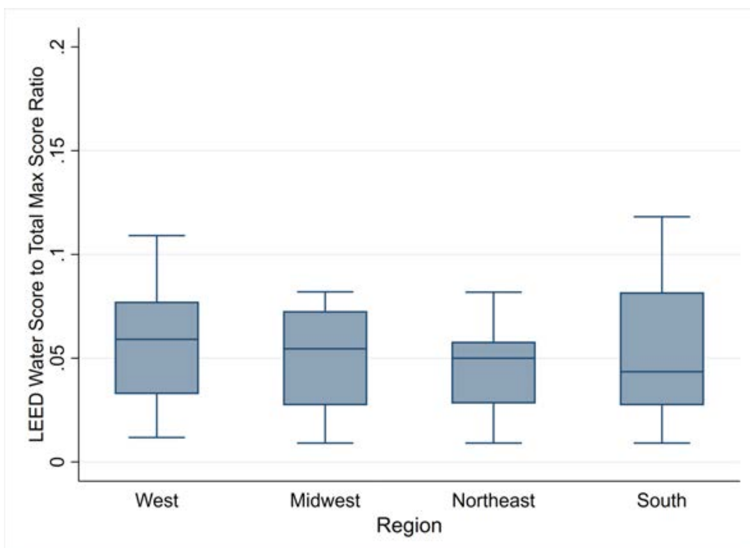
(d) Source EUI with 3NN

Notes: This figure provides robustness checks for the difference in differences approach. They repeat the exercise in Figure A.7 with the untreated buildings only. The untreated building sample is constructed following the procedures in Appendix III. The reference year is set to be one year prior to the LEED program registration date. (a) and (b) use the site energy use intensity as the dependent variable and report the regression coefficients after 5NN and 3NN propensity score matching. (c) and (d) use source energy use intensity as the outcome variable. The vertical bars around the coefficients indicate the 95% confidence interval.

Figure A.6: Energy and Water Component Score Distribution Across Census Regions



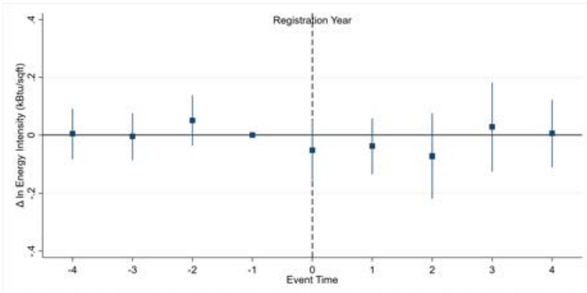
(a) Energy Component Score



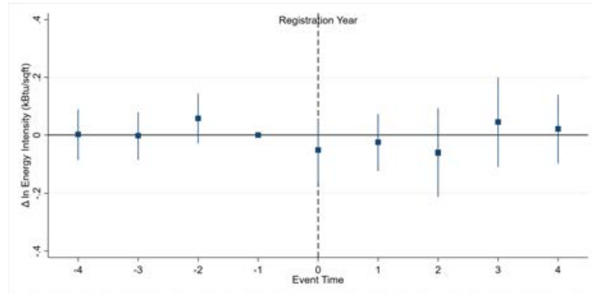
(b) Water Component Score

Notes: This figure provides energy and water component score distributions across the four census regions in the United States. Surprisingly, the score distributions are relatively similar across all four regions.

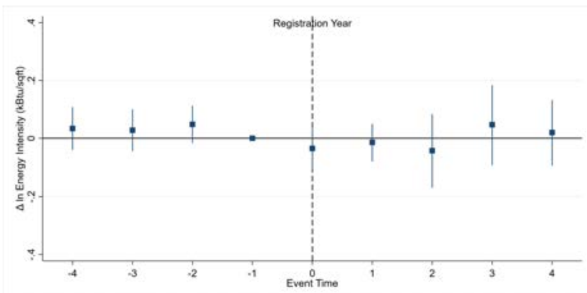
Figure A.7: Energy Savings in Each Period Relative to the Reference Year Savings
(Different Combinations of Energy Sources and Matching Procedures)



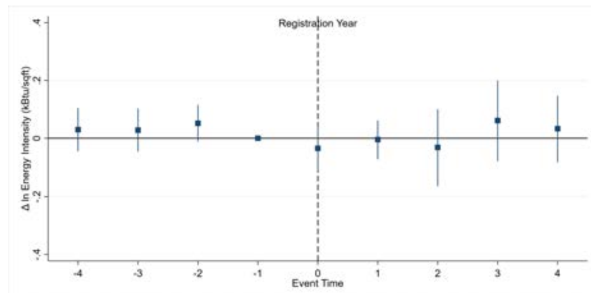
(a) Site EUI with 5NN



(b) Site EUI with 3NN



(c) Source EUI with 5NN



(d) Source EUI with 3NN

Notes: This figure plots the δ's from the event study regression in Equation (3). The reference year is set to be one year prior to the LEED program registration date. (a) and (b) use the site energy use intensity as the dependent variable and report the regression coefficients after the propensity score matching with 5NN and 3NN. (c) and (d) plot the coefficients with source energy use intensity as the outcome variable. The vertical bars around the coefficients indicate the 95% confidence interval.

Table A.1: Summary of Sample Construction Criteria
(Number of Unique Buildings)

	All Buildings	LEED Buildings
Federally Owned GSA Buildings	562	110
Data for 4 Years Before & After LEED Registration	-43	-43
Buildings with Complete LEED Component Scores	-7	-7
Number of Buildings in the Main Results	512	60

Notes: This table summarizes the changes in sample sizes with corresponding sample construction criteria. We have a total of 562 federally owned GSA buildings in the dataset. As discussed in Section 3, the LEED buildings are identified based on their address in the EUAS dataset and the GBIG reports. 7 out of 110 LEED buildings record vague/inaccurate street address in the EUAS dataset, thus we manually map them to LEED buildings in the GBIG reports through Google Maps. The data restriction ensures that the LEED buildings have data at least 4 years of data before and after the treatment. Among the 43 buildings that are deducted from the complete data to construct our sample for the analysis, 38 of them do not contain at least 4 years of data before the treatment date, where most of them only start recording their consumption after registration year. 1 building does not report enough post treatment data, and the remaining 4 buildings have missing data points within these periods. Due to different types of the LEED program, 7 buildings do not report the complete six components scores. The main results in the paper are estimated using the panel with complete LEED scores. This sample consists of 60 LEED buildings and 452 non-LEED buildings. 4 out of 60 LEED buildings in the sample were certified more than once, and we assign their first registration year as the treatment indicator in our analysis. For robustness check, estimates remain stable if we exclude those 4 buildings from the treatment group. Also, for the record, we exclude one LEED building from the dataset because it has an abnormal consumption trend over the years. We also exclude five non-LEED buildings from the dataset because they were certified with the Green Globes. Results with the inclusion of these buildings are similar to the main analysis.

Table A.2: Summary of Relevant LEED Programs in the Sample for Main Results

	Number of LEED-Certified Retrofitted Buildings
LEED NC 2.1	6
LEED NC 2.2	4
LEED NC 2009	9
LEED EB 2.0	1
LEED EB 2009	34
LEED CI 2.0	3
LEED CI 2009	1
LEED CS 2.0	1
LEED CS 2009	1
Total Frequency	60

Notes: This table presents the types of LEED programs involved in the main results estimation. Although multiple LEED programs are involved in the analysis, more than half of them are certified with LEED EB 2009. Glossary of the program abbreviations are: LEED NC (LEED for New Construction); LEED EB (LEED for Existing Buildings); LEED CI (LEED for Commercial Interiors); LEED CS (LEED for Core & Shell). It is important to note that existing buildings can obtain LEED certification as new construction because major renovation exceeding 60% of the entire square footage is considered new construction. As we can see in this table, a third of the LEED-certified retrofitted buildings in our sample obtained certification as *new construction*.

Table A.3: Correlation Matrix of Attributes in the LEED Program

	CR	ES	SS	IS	MS	WS	OS
CR	1.000						
ES	0.144	1.000					
SS	-0.012	0.174	1.000				
IS	-0.169	-0.428	0.089	1.000			
MS	-0.275	-0.429	0.148	0.674	1.000		
WS	-0.065	0.197	0.120	0.004	0.029	1.000	
OS	-0.238	-0.423	0.023	0.533	0.514	0.016	1.000

Notes: This table presents the correlation matrix for different attributes in the LEED program. CR stands for certified ratio, and refers to the fraction of a building square footage that is LEED-certified. ES, SS, IS, MS, WS, and OS represent the component scores evaluated by the LEED program: energy and atmosphere, sustainable sites, indoor environmental quality, materials and resources, water efficiency, and innovation in operation, respectively. As observed, MS, IS, and OS, which are relatively highly correlated.

Table A.4: Energy Use Intensity Impacts from the LEED Program and Heterogeneity Analysis with Energy Star Program

	All Buildings					Buildings W/O Energy Star
	(1)	(2)	(3)	(4)	(5)	(6)
T	-0.033	-0.030	0.026	-0.011	-0.019	0.018
	(0.045)	(0.044)	(0.062)	(0.055)	(0.055)	(0.063)
T × CR				-0.065**	-0.089**	
				(0.027)	(0.041)	
T × ES				-0.138***	-0.125***	
				(0.042)	(0.044)	
T × SS				0.016	0.013	
				(0.046)	(0.065)	
T × WS				0.056*	0.085*	
				(0.033)	(0.044)	
T × MIIS				-0.023	0.003	
				(0.030)	(0.044)	
T × Energy Star			-0.152**	-0.095	-0.103	
			(0.073)	(0.069)	(0.067)	
T × CR × Energy Star					0.114**	
					(0.057)	
T × ES × Energy Star				0.072	0.059	
				(0.068)	(0.059)	
T × SS × Energy Star					-0.052	
					(0.076)	
T × WS × Energy Star					-0.092	
					(0.058)	
T × MIIS × Energy Star					-0.043	
					(0.054)	
Energy Star		-0.088**	-0.059	-0.058	-0.059	
		(0.040)	(0.040)	(0.041)	(0.041)	
Year FE	×	×	×	×	×	×
Building FE	×	×	×	×	×	×
Adj. R-squared	0.665	0.667	0.668	0.671	0.672	0.756
Observations	9,599	9,599	9,599	9,599	9,599	6,069

Notes: This table reports energy savings from the LEED program in federal buildings and the heterogeneity treatment effects with Energy Star program. The dependent variable is the log of site energy use intensity, and the untreated group is constructed using five-nearest-neighbor (5NN) matching with replacement. T is the treatment indicator equal to one if a building i in year t is registered with LEED program, and zero otherwise. CR stands for certified ratio. ES, SS, WS, and MIIS represent the component scores evaluated by the LEED program: energy and atmosphere, sustainable sites, water efficiency, and the combination of material, indoor environment, and innovation, respectively. To facilitate comparison, CR, ES, SS, WS, and MIIS are all standardized to have mean zero and standard deviation one. Energy Star is the treatment indicator for Energy Star program equal to one if a building i was labeled in year t. Columns (1)-(5) include all federal buildings, and column (6) only includes LEED and non-LEED buildings without Energy Star labels. Efficiency Conservation Measures (ECM), CDD and HDD are included as controls in all regressions. Because of the potentially repeated non-LEED buildings in the untreated group due to matching with replacement, all regressions are weighted by the inverse of the number of repeated observations. Standard errors clustered by building are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A.5: Energy Use Intensity Impacts from Different Combinations of the Energy and Water Component Scores

	1 S.D. Decrease in ES	1 S.D. Increase in ES
Panel A: All LEED Buildings		
1 S.D. Decrease in WS	0.025	-0.227
1 S.D. Increase in WS	0.151	-0.101
Panel B: LEED Office Buildings		
1 S.D. Decrease in WS	0.057	-0.221
1 S.D. Increase in WS	0.223	-0.055

Notes: This table shows the energy use intensity impacts with different combinations of the energy and water component scores. The values are calculated based on columns (2) and (3) from Table 4 – for all buildings and office buildings, respectively – and they provide an easier way to understand those estimated coefficients. ES and WS in the row and column labels represent component scores of energy and atmosphere and water efficiency, respectively. Each number in this table represents the average percentage change in building energy use when increasing or decreasing one standard deviation (S.D.) of the corresponding score component. For example, at the extremes, a building that had a one standard deviation higher energy score than average and one standard deviation lower water score than average would have energy use that was 22.7% lower than a non-LEED building.

Table A.6: Water Use Intensity Outcome and Heterogeneity Analysis

	(1) All Buildings	(2) Office	(3) All Buildings	(4) Office
T	-0.040	-0.056	-0.083	-0.089
	(0.056)	(0.070)	(0.062)	(0.070)
T × CR			0.081**	0.084**
			(0.040)	(0.040)
T × ES			0.083*	0.093*
			(0.045)	(0.055)
T × SS			0.095*	0.087
			(0.048)	(0.058)
T × WS			0.064	0.066
			(0.041)	(0.053)
T × MIIS			-0.099	-0.095
			(0.104)	(0.112)
Year FE	×	×	×	×
Building FE	×	×	×	×
Adj. R-squared	0.753	0.719	0.756	0.723
Observations	4,985	3,510	4,985	3,510

Notes: This table reports water savings from the LEED program in federal buildings and the heterogeneous treatment effects regarding different program dimensions when restricting the sample to buildings that reported water consumption data. The dependent variable is the log of site water use intensity in the corresponding columns. The untreated group is constructed using five-nearest-neighbor (5NN) matching with replacement. T is the treatment indicator equal to one if a building i in year t is registered with LEED, and zero otherwise. CR stands for certified ratio, and refers to the fraction of a building square footage that is LEED-certified. ES, SS, WS, and MIIS represent the component scores evaluated by the LEED program: energy and atmosphere, sustainable sites, water efficiency, and the combination of material, indoor environment, and innovation, respectively. To facilitate comparison, CR, ES, SS, WS, and MIIS are all standardized to have mean zero and standard deviation one. Columns (1) and (3) include all federal buildings, and columns (2) and (4) include office buildings only. CDD and HDD are included as controls in all regressions. Because of the potentially repeated non-LEED buildings in the untreated group due to matching with replacement, all regressions are weighted by the inverse of the number of repeated observations. Standard errors clustered by building are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A.7: Energy Use Intensity Outcome and Heterogeneity Analysis
(Cooling Degree Days (CDD) and Heating Degree Days (HDD))

	(1)	(2)
T	-0.030	-0.014
	(0.045)	(0.178)
CDD	0.000*	0.000*
	(0.000)	(0.000)
HDD	0.000***	0.000***
	(0.000)	(0.000)
T × CDD		0.000
		(0.000)
T × HDD		-0.000
		(0.000)
Year FE	×	×
Building FE	×	×
Adj. R-squared	0.665	0.665
Observations	9,599	9,599

Notes: This table reports energy savings from the LEED program in federal buildings and the heterogeneity treatment effects regarding Cooling Degree Days (CDD) and Heating Degree Days (HDD). The dependent variable is the log of site water use intensity in the corresponding columns. The untreated group is constructed using five-nearest-neighbor (5NN) matching with replacement. T is the treatment indicator equal to one if a building i in year t is registered with LEED, and zero otherwise. Columns (1) and (2) include all federal buildings. Because of the potentially repeated non-LEED buildings in the untreated group due to matching with replacement, all regressions are weighted by the inverse of the number of repeated observations. Standard errors clustered by building are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A.8: Federal Space Standards, Criteria, and Guidelines

Grade	Type of assignment 1[2]	Office space[2]
GS 1 to 6		60
GS 7 to 11	Nonsupervisory	75
GS 7 to 11	Supervisory	100
GS 12 to 13	Nonsupervisory	100
GS 12 to 13	Supervisory	150
GS 14 to 15	Nonsupervisory	150
GS 14 to 15	Supervisory	225
GS 16, 17, and 18	Nonsupervisory	225
GS 16	Supervisory	300
GS 17	Supervisory	350
GS 18	Supervisory	400

Notes: [1] Supervisory means supervision of or frequent meetings with 3 or more employees within the office's confines; [2] Allowance in square feet per person;

This table presents the spece standards, criteria, and guidelines for federal buildings. It is obtained from `https://www.govinfo.gov/content/pkg/CFR-1997-title41-vol2/xml/CFR-1997-title41-vol2-part101-id360-subpart101-id430.xml`.

Table A.9: Energy Use Intensity Outcome and Heterogeneity Analysis
(Different Combinations of Energy Sources and Matching Methods)

	Site Energy				Source Energy			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	5NN	5NN	3NN	3NN	5NN	5NN	3NN	3NN
T	-0.038	0.001	-0.012	0.015	-0.049	-0.017	-0.010	-0.007
	(0.040)	(0.051)	(0.049)	(0.062)	(0.036)	(0.045)	(0.044)	(0.056)
T × CR	-0.054*	-0.041	-0.018	0.001	-0.056*	-0.043	-0.027	-0.010
	(0.028)	(0.030)	(0.043)	(0.046)	(0.028)	(0.028)	(0.042)	(0.040)
T × ES	-0.126***	-0.139***	-0.177***	-0.193***	-0.110***	-0.105**	-0.162***	-0.150***
	(0.037)	(0.048)	(0.045)	(0.059)	(0.035)	(0.044)	(0.041)	(0.052)
T × SS	0.024	0.038	0.011	0.032	0.023	0.031	0.022	0.035
	(0.046)	(0.051)	(0.050)	(0.055)	(0.034)	(0.038)	(0.042)	(0.046)
T × WS	0.063*	0.083**	0.069*	0.106**	0.052*	0.074**	0.057*	0.096**
	(0.033)	(0.035)	(0.036)	(0.041)	(0.031)	(0.032)	(0.034)	(0.038)
T × MIIS	-0.008	-0.015	-0.000	-0.014	0.002	0.004	-0.002	-0.011
	(0.032)	(0.037)	(0.046)	(0.049)	(0.031)	(0.036)	(0.043)	(0.048)
Year FE	×	×	×	×	×	×	×	×
Building FE	×	×	×	×	×	×	×	×
Adj. R-squared	0.669	0.565	0.699	0.597	0.702	0.622	0.736	0.657
Observations	9,599	6,894	6,483	4,600	9,599	6,894	6,483	4,600

Notes: This table reports energy savings from the LEED program in federal buildings and the heterogeneity treatment effects regarding different program dimensions with different matching method and energy sources. The dependent variable is the log of site energy use intensity, and the untreated group is constructed using five-nearest-neighbor (5NN) and three-nearest-neighbor (3NN) matching with replacement. T is the treatment indicator equal to one if a building i in year t is registered with LEED program, and zero otherwise. CR stands for certified ratio, and refers to the fraction of a building square footage that is LEED-certified. ES, SS, WS, and MIIS represent the component scores evaluated by the LEED program: energy and atmosphere, sustainable sites, water efficiency, and the combination of material, indoor environment, and innovation, respectively. To facilitate comparison, CR, ES, SS, WS, and MIIS are all standardized to have mean zero and standard deviation one. Columns (1), (3), (5), and (7) include all federal buildings, and columns (2), (4), (6), and (8) include office buildings only. As observed, the coefficients are consistent with the results reported in Table 4 columns (7) and (8). CDD and HDD are included as controls in all regressions. Because of the potentially repeated non-LEED buildings in the untreated group due to matching with replacement, all regressions are weighted by the inverse of the number of repeated observations. Standard errors clustered by building are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A.10: Energy Use Intensity Outcome and Heterogeneity Analysis with Different Specification and Propensity Score Matching Strategy

	Matching W/ Year 2000		Matching W/O Repeated Buildings		Matching W/ SD		Matching Within Building Type	
	(1) All Buildings	(2) Office	(3) All Buildings	(4) Office	(5) All Buildings	(6) Office	(7) All Buildings	(8) Office
T	-0.014	0.028	-0.038	0.001	-0.038	0.002	-0.021	-0.020
	(0.042)	(0.054)	(0.040)	(0.051)	(0.040)	(0.048)	(0.039)	(0.042)
T × CR	-0.064**	-0.049	-0.054*	-0.041	-0.059**	-0.043	-0.051*	-0.053*
	(0.032)	(0.036)	(0.028)	(0.030)	(0.028)	(0.031)	(0.029)	(0.028)
T × ES	-0.138***	-0.148***	-0.126***	-0.139***	-0.130***	-0.140***	-0.111***	-0.122***
	(0.036)	(0.047)	(0.037)	(0.048)	(0.038)	(0.048)	(0.039)	(0.043)
T × SS	0.006	0.018	0.024	0.038	0.028	0.040	-0.011	-0.005
	(0.050)	(0.055)	(0.046)	(0.052)	(0.047)	(0.052)	(0.033)	(0.038)
T × WS	0.073**	0.075**	0.063*	0.083**	0.072**	0.082**	0.057	0.087**
	(0.034)	(0.037)	(0.033)	(0.035)	(0.032)	(0.035)	(0.034)	(0.035)
T × MIIS	-0.024	-0.028	-0.008	-0.015	-0.001	-0.013	0.000	-0.002
	(0.034)	(0.039)	(0.032)	(0.037)	(0.031)	(0.036)	(0.031)	(0.033)
Year FE	×	×	×	×	×	×	×	×
Building FE	×	×	×	×	×	×	×	×
Adj. R-squared	0.614	0.603	0.663	0.556	0.732	0.675	0.670	0.693
# LEED Buildings	52	41	60	45	58	45	54	44
Observations	8,954	6,814	5,628	3,978	9,033	6,361	8,629	7,246

Notes: This table reports results from several robustness checks by re-estimating the specifications of columns (2) and (3) of Table 4 with different first stage propensity score matching strategy. T is the treatment indicator equal to one if a building i in year t is registered with LEED, and zero otherwise. CR stands for certified ratio, and refers to the fraction of a building square footage that is LEED-certified. ES, SS, WS, and MIIS represent the component scores evaluated by the LEED program: energy and atmosphere, sustainable sites, water efficiency, and the combination of material, indoor environment, and innovation, respectively. To facilitate comparison, CR, ES, SS, WS, and MIIS are all standardized to have mean zero and standard deviation one. The dependent variable is the log of site energy use intensity. The untreated group is constructed using five-nearest-neighbor (5NN) matching with replacement. Columns (1) and (2) use the energy use intensity in year 2000 instead of the pre-treatment energy use intensity for the propensity score matching to construct the sub-sample for second stage regressions. Columns (3) and (4) drop the repeated non-LEED buildings in the untreated group due to matching with replacement. Columns (5) and (6) add the standard deviation of log site energy use intensity in the pre-treatment periods in the propensity score matching specification. Columns (7) and (8) match LEED buildings to the same types of non-LEED buildings. Columns (1), (3), (5), and (7) include all federal buildings, and columns (2), (4), (6), and (8) include office buildings only. CDD and HDD are included as controls in all regressions. Because of the potentially repeated non-LEED buildings in the untreated group due to matching with replacement, all regressions except for columns (3) and (4) are weighted by the inverse of the number of repeated observations. Note columns (3) and (4) remove the repeated non-LEED buildings from the sample, and thus they are not weighted. Standard errors clustered by building are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A.11: Energy Use Intensity Outcome and Heterogeneity Analysis with and without LEED New Construction Certification

	Main Results	LEED NC Discussion		Time Intervals B/W Registration and Awarded Years	
	(1)	(2) W/O LEED NC	(3) W/LEED NC Only	(4) Shorter Time Interval	(5) Longer Time Interval
T	-0.038	-0.039	-0.001	-0.084**	0.079
	(0.040)	(0.044)	(0.073)	(0.040)	(0.080)
T × CR	-0.054*	-0.006	-0.138**	-0.036*	-0.069
	(0.028)	(0.030)	(0.060)	(0.020)	(0.053)
T × ES	-0.126***	-0.080**	-0.158**	-0.075***	-0.176*
	(0.037)	(0.034)	(0.068)	(0.022)	(0.096)
T × SS	0.024	0.065	-0.027	-0.011	0.054
	(0.046)	(0.061)	(0.084)	(0.029)	(0.103)
T × WS	0.063*	0.072**	0.082	0.071**	0.075
	(0.033)	(0.034)	(0.079)	(0.034)	(0.097)
T × MIIS	-0.008	0.013	0.042	-0.052*	-0.004
	(0.032)	(0.046)	(0.060)	(0.029)	(0.067)
Year FE	×	×	×	×	×
Building FE	×	×	×	×	×
Adj. R-squared	0.669	0.697	0.510	0.539	0.695
Observations	9,599	6,551	3,048	6,020	3,579

Notes: This table reports results for LEED buildings with and without LEED New Construction (NC) Certification and LEED buildings that were certified shortly or later after program registration, as robustness checks by re-estimating the specifications of columns (2) and (3) of Table 4. T is the treatment indicator equal to one if a building i in year t is registered with LEED, and zero otherwise. CR stands for certified ratio, and refers to the fraction of a building square footage that is LEED-certified. ES, SS, WS, and MIIS represent the component scores evaluated by the LEED program: energy and atmosphere, sustainable sites, water efficiency, and the combination of material, indoor environment, and innovation, respectively. To facilitate comparison, CR, ES, SS, WS, and MIIS are all standardized to have mean zero and standard deviation one. The dependent variable is the log of site energy use intensity. The untreated group is constructed using five-nearest-neighbor (5NN) matching with replacement. Column (1) shows the estimates of our preferred specification in Table 4 column 2. Column (2) uses LEED buildings that were certified other than NC and column (3) uses LEED buildings that were certified with LEED NC. Column (4) includes LEED buildings that were certified within three years of registration, and column (5) includes LEED buildings that were certified more than three years from registration. Note that the average time interval between registration and awarded years in our dataset is three years. All columns include all federal buildings. Comparing to the main results in Table 4, the estimates for different attributes of the LEED program are similar. CDD and HDD are included as controls in all regressions. Because of the potentially repeated non-LEED buildings in the untreated group due to matching with replacement, both regressions are weighted by the inverse of the number of repeated observations. Standard errors clustered by building are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A.12: Energy Use Intensity Outcome and Heterogeneity Analysis with Different Specification and Sample Restrictions

	W/ High Certified Ratio		W/ Data Points More than Six Months in a Year		W/ More Post Data		Multiple Weights	
	(1) All Buildings	(2) Office	(3) All Buildings	(4) Office	(5) All Buildings	(6) Office	(7) All Buildings	(8) Office
T	-0.011	0.045	-0.028	0.007	-0.025	0.002	-0.064*	-0.054
	(0.059)	(0.071)	(0.040)	(0.050)	(0.040)	(0.050)	(0.035)	(0.039)
T × CR	-0.132	-0.144	-0.056**	-0.042	-0.059**	-0.048	-0.023	-0.023
	(0.098)	(0.106)	(0.028)	(0.030)	(0.029)	(0.032)	(0.023)	(0.021)
T × ES	-0.132***	-0.154***	-0.128***	-0.139***	-0.134***	-0.147***	-0.083**	-0.103**
	(0.032)	(0.043)	(0.035)	(0.046)	(0.034)	(0.046)	(0.041)	(0.041)
T × SS	0.027	0.051	0.013	0.025	0.011	0.026	0.010	0.023
	(0.055)	(0.062)	(0.046)	(0.051)	(0.046)	(0.051)	(0.025)	(0.026)
T × WS	0.055	0.066	0.057*	0.075**	0.052	0.070**	0.043	0.084***
	(0.038)	(0.041)	(0.032)	(0.034)	(0.032)	(0.034)	(0.030)	(0.026)
T × MIIS	-0.027	-0.046	-0.012	-0.017	-0.014	-0.020	0.006	-0.002
	(0.040)	(0.047)	(0.030)	(0.035)	(0.030)	(0.034)	(0.029)	(0.030)
Year FE	×	×	×	×	×	×	×	×
Building FE	×	×	×	×	×	×	×	×
Adj. R-squared	0.666	0.556	0.708	0.596	0.718	0.636	0.580	0.547
Observations	9,391	6,714	9,422	6,789	9,545	6,816	9,599	6,894

Notes: This table reports results from several robustness checks by re-estimating the specifications of columns (2) and (3) of Table 4 with different restrictions. T is the treatment indicator equal to one if a building i in year t is registered with LEED, and zero otherwise. CR stands for certified ratio, and refers to the fraction of a building square footage that is LEED-certified. ES, SS, WS, and MIIS represent the component scores evaluated by the LEED program: energy and atmosphere, sustainable sites, water efficiency, and the combination of material, indoor environment, and innovation, respectively. To facilitate comparison, CR, ES, SS, WS, and MIIS are all standardized to have mean zero and standard deviation one. The dependent variable is the log of site energy use intensity. The untreated group is constructed using five-nearest-neighbor (5NN) matching with replacement. Columns (1) and (2) only include buildings with certified ratio greater than 50%. Columns (3) and (4) restrict the sample to observations with at least six months per year when annualized. Columns (5) and (6) restrict the LEED buildings to have at least six years of post-certification data instead of four years as in the main sample. Columns (7) and (8) estimate the main specifications with multiple weights. Columns (1), (3), (5), and (7) include all federal buildings, and columns (2), (4), (6), and (8) include office buildings only. Comparing to the main results in Table 4, the estimates for different attributes of the LEED program are similar. CDD and HDD are included as controls in all regressions. Because of the potentially repeated non-LEED buildings in the untreated group due to matching with replacement, columns (1)-(6) are weighted by the inverse of the number of repeated observations. Columns (7) and (8) are weighted by both the inverse of the number of repeated observations and the building sizes. Standard errors clustered by building are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A.13: Energy Use Intensity Outcome and Heterogeneity Analysis
(Different Year Restrictions)

	1990 - 2019		1995 - 2019		2000 - 2019	
	(1) All Buildings	(2) Office	(3) All Buildings	(4) Office	(5) All Buildings	(6) Office
T	-0.038	0.001	-0.041	-0.001	-0.045	-0.007
	(0.040)	(0.051)	(0.038)	(0.047)	(0.036)	(0.044)
T × CR	-0.054*	-0.041	-0.042*	-0.029	-0.032	-0.022
	(0.028)	(0.030)	(0.025)	(0.027)	(0.023)	(0.024)
T × ES	-0.126***	-0.139***	-0.121***	-0.141***	-0.084**	-0.104**
	(0.037)	(0.048)	(0.035)	(0.044)	(0.035)	(0.041)
T × SS	0.024	0.038	0.023	0.039	0.016	0.033
	(0.046)	(0.051)	(0.042)	(0.047)	(0.041)	(0.044)
T × WS	0.063*	0.083**	0.055*	0.076**	0.031	0.049
	(0.033)	(0.035)	(0.031)	(0.033)	(0.030)	(0.032)
T × MIIS	-0.008	-0.015	-0.006	-0.019	0.003	-0.015
	(0.032)	(0.037)	(0.032)	(0.036)	(0.032)	(0.036)
Year FE	×	×	×	×	×	×
Building FE	×	×	×	×	×	×
Adj. R-squared	0.669	0.565	0.698	0.606	0.740	0.663
Observations	9,599	6,894	8,346	5,936	6,955	4,915

Notes: This table reports energy savings from the LEED program in federal buildings and the heterogeneous treatment effects regarding different program dimensions when restricting the sample to different time periods as specified in the column titles. The dependent variable is the log of site energy use intensity, and the untreated group is constructed using five-nearest-neighbor (5NN) matching with replacement. T is the treatment indicator equal to one if a building i in year t is registered with LEED, and zero otherwise. CR stands for certified ratio, and refers to the fraction of a building square footage that is LEED-certified. ES, SS, WS, and MIIS represent the component scores evaluated by the LEED program: energy and atmosphere, sustainable sites, water efficiency, and the combination of material, indoor environment, and innovation, respectively. To facilitate comparison, CR, ES, SS, WS, and MIIS are all standardized to have mean zero and standard deviation one. Columns (1), (3), and (5) include all federal buildings, and columns (2), (4), and (6) include office buildings only. CDD and HDD are included as controls in all regressions. Because of the potentially repeated non-LEED buildings in the untreated group due to matching with replacement, all regressions are weighted by the inverse of the number of repeated observations. Standard errors clustered by building are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.