Estimating the effect of energy disclosure requirements on building energy use

Ben Hellerstein, Paul H. O'Neill School of Public and Environmental Affairs, Indiana University

ABSTRACT

In recent years, dozens of U.S. cities and states have adopted policies requiring the owners of large buildings to disclose the amount of energy used in their buildings on an annual basis. Building Energy Disclosure and Benchmarking (BEDB) policies may reduce energy consumption by encouraging owners to invest in energy efficiency, but estimating the effect of BEDB on energy use is challenging because data availability is closely tied to the treatment. In other words, the energy use data sets generated by BEDB policies include only the buildings required to report their energy consumption, omitting all buildings that are not subject to the policy. Taking advantage of changes in the square-footage threshold for buildings that are required to comply with BEDB policies in 11 cities, this paper estimates the effect of disclosure requirements on energy use. Using a regression discontinuity design, I estimate the local treatment effect in the vicinity of the compliance threshold for commercial buildings that have been exposed to several years of BEDB requirements (treatment group), relative to commercial buildings experiencing their first or second year of BEDB requirements (control group). The analysis finds a statistically significant difference between the treatment and control groups, although the magnitude of the estimated effect may be implausibly large, suggesting that other complementary policies are introducing bias. With further refinements, the model presented in this paper is a promising approach to generate robust causal estimates of the effect of BEDB and other information disclosure policies on energy and environmental outcomes.

1. Introduction

Reducing energy consumption in buildings is essential to meeting local and international greenhouse gas emissions targets, but the pace of energy efficiency improvements is far below what is needed (Nadel and Hinge 2020). The "energy efficiency gap" refers to the difference between the level of energy efficiency improvements that would minimize social costs and the level that is implemented. Academics and private-sector analysts disagree about the size of the gap, particularly related to the unrealized potential for private cost savings, but there appear to be some investment inefficiencies that could be addressed through well-designed policy (Allcott and Greenstone 2012).

One increasingly popular tool to help close this gap is to require the owners of large commercial and residential buildings to report the amount of energy consumed in their buildings on an annual basis, a policy known as Building Energy Disclosure and Benchmarking (BEDB). Typically, each building is assigned a score based on its energy performance relative to similar buildings, and the energy consumption data is made available to the public. Most BEDB policies set a minimum threshold for the size of buildings that are required to report their energy use, expressed in terms of square footage or number of residential units.

BEDB policies could reduce energy use through several mechanisms. A BEDB policy may enable owners to compare their building's energy consumption with buildings of a similar size and use, and identify opportunities for efficiency improvements that reduce the owner's operating costs (Krukowski and Burr 2012; Mims et al. 2017; Palmer and Walls 2015a). BEDB policies also may allow prospective renters or buyers to compare energy use among multiple buildings. Over time, as consumers prefer more energy-efficient buildings, owners of those buildings will realize a lease or sale premium, incentivizing owners to invest in energy efficiency improvements (Mims et al. 2017; Palmer and Walls 2015a; Allcott and Greenstone 2012).

Assessing the effect of BEDB policies on energy use is complicated by the fact that the availability of data is closely tied to the treatment. The datasets collected by city and state governments under these policies generally include only those buildings that are required to report their energy use. Datasets allowing for direct comparisons between treated and untreated buildings are harder to find. The few causal studies on BEDB that have been conducted use difference-in-difference identification strategies relying on the construction of an appropriate control group (Meng, Hsu, and Han 2017; Palmer and Walls 2015b), or in one case a randomized controlled field experiment featuring a small sample size and a limited range of building types (Woodson et al. 2015).

This paper examines the effect of BEDB policies in 11 cities in the United States. I use a regression discontinuity design (RDD) to identify the effect of BEDB on energy use for buildings near the square-foot threshold for compliance. If the assumptions of RDD are met, buildings can be assumed to be assigned to treatment and control groups on an as-good-as-random basis, and the estimated difference between the groups at the treatment cutoff is plausibly causal. The analysis relies on staggered adoption timeline in cities where the square-footage threshold for compliance was lowered over time to estimate the difference in energy use between a treatment group consisting of commercial buildings that have been exposed to multiple years of BEDB, and a control group of buildings that are in their first or second year of required compliance with BEDB.

The preliminary analysis shows a substantial difference in total site energy use (TSE), the quantity of energy consumed in a building, between the control and treatment group in the vicinity of the compliance threshold. In the full data set, including all types of commercial buildings across the 11 cities in the study, TSE near the threshold is 19.4–20.9% lower in buildings that have been subject to BEDB policies for multiple years, compared to buildings that are in their first or second year of compliance. When the data set is restricted only to office buildings, energy use is 13.2% lower in the treatment group compared to the control group. The estimated treatment effect in the models including all commercial buildings is statistically significant, and the magnitude of the effect is substantially higher than other studies on BEDB have found. Further refinements to the analysis will consider the presence of other policies with similar compliance thresholds that may introduce bias into these estimates.

While the results in this paper are preliminary and should not be taken as a definitive estimate of the magnitude of the effect of BEDB policies, the analysis introduces a promising method to evaluate the effectiveness of BEDB and other policies aiming to change energy and environmental outcomes through the disclosure of information. This paper introduces three innovations into the study of information disclosure policies: 1) aggregating data from nearly a dozen jurisdictions to increase statistical power and generate more precise estimates; 2) overcoming a lack of data on untreated buildings by using data from cities where compliance thresholds that have been lowered over time, allowing for comparisons between buildings that have been exposed to a greater versus lesser duration of treatment; and 3) using RDD to increase confidence in the causal interpretation of the results. With additional development, this approach will generate robust estimates of the effect of BEDB on energy consumption, providing greater certainty to public officials, advocates, and industry professionals designing strategies to reduce building sector energy use.

2. Information disclosure and energy efficiency

2.1. Building energy disclosure policies

As of January 2026, 8 states and 50 U.S. cities have passed BEDB policies, with the majority of those policies having gone into effect in the last 10 years (Institute for Market Transformation 2026). BEDB policies are part of a broader trend toward transparency and information disclosure requirements in environmental policy and other policy areas (Stavins 2003; Weil et al. 2006).

BEDB policies typically require the disclosure of all energy consumed in a building, regardless of who pays for it. In some multi-tenant buildings, tenants pay for all utility costs associated with energy consumption in the space they rent. In other buildings, utilities are master-metered, meaning that the amount of energy consumed is not tracked on a per-tenant basis. The landlord pays the utility bills, and the cost is usually incorporated into the rent charged to tenants. In buildings where tenant spaces are separately metered for electricity, other utilities like hot water and space heating may be master-metered. Additionally, owners typically pay for the energy used in common spaces like entries and hallways.

Public officials adopt BEDB policies with several goals in mind. The data collected under BEDB may help inform the future direction of public investments in energy efficiency and climate mitigation programs (Beddingfield, Hughes, and Hart 2017). In some cases, a BEDB policy may be the first step toward adopting building performance standards (BPS) requiring building owners to reduce the amount of energy their buildings consume.

Establishing a BEDB policy may drive improvements in energy efficiency even if other policies and investments remain constant. I separate the mechanisms by which this energy reduction can happen into two categories, following Mims et al. (2017) and Meng, Hsu, and Han (2017).

The first type of mechanism involves actions taken by the building owner to reduce operating costs in direct response to the availability of energy use data. These actions are motivated by the fact that the owner pays some portion of the utility costs in buildings where utilities are separately metered (for energy consumed in public areas of the building like hallways and lobbies), and up to 100% of the utility costs in master-metered buildings. The second type of mechanism is mediated by the preferences of potential renters and buyers who seek to reduce their utility costs in buildings where tenants pay their own utility bills. Tenants may use data disclosed through BEDB policies to choose more energy-efficient buildings, driving up the price of those properties and incentivizing owners of less-efficient properties to make improvements. The next two sections describe the evidence for each type of mechanism.

2.2. Mechanisms operating directly on the building owner

A building owner has a direct incentive to invest in efficiency improvements that reduce master-metered energy consumption in tenant spaces, as well as energy used in public areas of the building. While owners already have information about their own buildings' energy consumption, BEDB may provide new information on energy use in other buildings of a similar size and use. Experiments in the single-family residential sector found that providing homeowners with data on how their energy consumption compares to their peers led to a decrease in energy use (Ayres, Raseman, and Shih 2009; Costa and Kahn 2010), although it is not known to what degree these findings translate to commercial or multifamily buildings.

In addition to providing novel information, BEDB may affect owners' decisions to invest in energy efficiency upgrades by increasing their attentiveness to the potential benefits. Individuals may make sub-optimal decisions when they fail to adequately consider all aspects of a choice problem, even when relevant information about costs and benefits is available to them (Allcott and Greenstone 2012). In some cases this inattention may be "rational," in the sense that the effort required to obtain or analyze the necessary information is too great (Palmer and Walls 2015a). By requiring owners to collect data on energy use and making it easy to compare energy performance across buildings, BEDB may make owners more attentive to energy costs while reducing the burden of acquiring information.

2.3. Mechanisms operating through consumer preferences

In buildings that are not master-metered, utility costs are part of the total cost for residents or commercial tenants to occupy a space. A tenant choosing between two properties that are equivalent in all aspects except for their level of energy consumption will therefore tend to choose the more efficient property if information about each property's energy performance is available to them.

Prior to the adoption of a BEDB policy, tenants know the energy consumption of the space they are currently occupying, but they may have little information about the energy performance of other buildings available for rent. By making energy performance data publicly available, BEDB allows consumers to express a preference for more efficient buildings. Over time, this will drive up the price of renting or buying a more efficient building, rewarding the owners of highly efficient buildings and creating an incentive for the owners of less efficient buildings to make improvements. As consumer preferences shift toward more efficient buildings, owners of those buildings may also see lower vacancy rates, leading to lower holding costs for investors (Fuerst and McAllister 2011).

There is evidence for a correlation between the energy efficiency of a building and its sale or lease price in situations where information about energy performance is available. Studies in Minnesota (Gilmer 1989; Laquatra 1986), the Australian Capital Territory (Department of the Environment, Water, Heritage and the Arts 2008), and the Netherlands (Brounen and Kok 2011) have shown a relationship between housing price and energy efficiency, controlling for other relevant factors, when home purchasers have access to information about how efficient a house is. Researchers have also found a relationship between LEED or Energy Star certification (two popular programs for sustainability certification) and commercial rents, using carefully constructed samples of non-certified buildings in the same neighborhood or sub-market as controls (Fuerst and McAllister 2011; Miller, Spivey, and Florance 2008).

2.4. Prior research on BEDB effectiveness

Many jurisdictions that have adopted BEDB policies report summary statistics and trends in energy use for buildings covered by the policy. These findings must be considered observational rather than causal, since they lack a control group or as-good-as-random assignment to treatment. Mims et al. (2017) synthesizes the results from many of the analyses, showing energy reductions that typically range between 3–8 percent over the course of 2–4 years.

Three studies have used causal research designs to generate credible estimates of the effect of BEDB policies on building energy use. Palmer and Walls (2015b) employed a difference-in-differences model to examine the early-adopter cities of Austin, New York, San Francisco, and Seattle, which implemented BEDB policies between October 2011 – June 2013. Using utility expense data from the National Council of Real Estate Investment Fiduciaries (NCREIF) between the first quarter of 2003 and the third quarter of 2013, they compared buildings in cities with BEDB policies to various control groups of buildings in jurisdictions that have not yet adopted BEDB. In their preferred specification, they found that utility bills decreased by 3% in covered buildings in cities with BEDB policies compared to a national control group. A control group restricted to the same metropolitan areas as the treated cities yielded a similar result, while a control group limited to other cities that adopted or considered adopting BEDB policies after the study period resulted in a statistically insignificant effect.

Meng, Hsu, and Han (2017) examined the BEDB program in New York City. The authors used a one-year difference in implementation timelines to set up a difference-in-differences analysis comparing privately owned office buildings as the treatment group and publicly owned office buildings as the control group. They considered the first year of reporting for privately owned buildings to be pre-treatment, since the owners are compiling and reporting the data for the previous year, when no energy data for their buildings had yet been publicly disclosed. They found a statistically insignificant increase in energy use in year 1 of treatment, a 6% reduction in energy use in year 2, and a 14.3% reduction in year 3. Meng, Hsu, and Han also found that disclosing Energy Star scores along with raw energy data had a greater impact on energy use than disclosing the raw data alone.

Woodson et al. (2015) reported the results of a randomized controlled experiment involving 500 multifamily residential buildings in Minnesota between 2012–2015. Buildings in the treatment group received a benchmarking service, which included analysis of utility data, access to an online platform, and support from an account manager. Master-metered buildings experienced a 5% decrease in energy use. Heating accounted for 82% of the observed savings in those buildings, and reductions in energy use were greater after the second year compared to the first year. The difference in energy use in non-master-metered buildings was not statistically significant, perhaps because the study was underpowered.

3. Methods

3.1. Regression discontinuity design

This paper uses a regression discontinuity design (RDD) to estimate the effect of BEDB on energy use in the vicinity of the threshold for policy compliance. An analysis based on RDD offers a key advantage over the difference-in-differences (DID) analyses used by Palmer and Walls (2015b) and Meng, Hsu, and Han (2017). DID requires researchers to designate a control group to serve as the counterfactual for what would have happened in the treated group if treatment had not been administered. To estimate causal effects using DID, researchers must assume that the treatment and control groups would have followed parallel trends in the absence of treatment. DID controls for time-invariant differences between treatment and control groups, but if the two groups differ in ways that vary with time, the estimated causal effect will be biased.

The risk of a biased DID estimate can be mitigated by demonstrating that treatment and control groups exhibited parallel trends prior to the beginning of treatment, and by showing that the results are robust to multiple specifications of the control group. Palmer and Walls (2015b) are able to show parallel trends in the pre-treatment period, and they obtain similar results with one alternative specification of the control group but not with the second alternative specification. Meng, Hsu, and Han (2017) are unable to examine pre-treatment trends because they only have data during one pre-treatment time period, and each of their analyses uses only one control group.

Even where parallel trends exist pre-treatment and the results are robust to multiple specifications of the control group, researchers cannot guarantee that parallel trends would have continued after treatment. Thus, the designation of a control group in DID analyses remains somewhat arbitrary.

Compared to DID, the RDD method provides a data-driven way to designate a control group. RDD takes advantage of discontinuities in the probability of treatment across the value of a certain variable. In this study, the variable determining assignment to treatment is the floor area of a building. At or above a certain square-footage threshold, building owners are required to comply with the BEDB policy; below the threshold, buildings are exempt from the policy. This is an example of a sharp RDD, since the probability of treatment is either 0 or 1 depending on which side of the cutoff a building falls on. The outcome variable is the site energy use intensity. This study investigates the difference in mean site energy use intensity between treated and untreated buildings at the cutoff.

RDD studies are sensitive to the choice of functional form to model the treatment and control group outcomes, as well as the bandwidth around the cutoff used for estimation. In this paper, I use methods described by Calonico, Cattaneo, and Farrell (2020) and implemented in the *rdrobust* R package (Calonico, Cattaneo, and Titiunik 2015) to choose the bandwidth that minimizes mean squared error in the estimation of the regression discontinuity, while allowing for robust inference that reduces the bias associated with estimating a regression model at a boundary.

RDD is a data-intensive statistical method, since only a portion of the observations in a data set (those near the treatment cutoff) are used for estimating the treatment effect. No single jurisdiction that adopted BEDB had data on enough buildings to provide sufficient statistical power for RDD. Thus, it was necessary to aggregate data across multiple jurisdictions. The following sections describe the sources of data and the process for preparing the data for use in RDD.

3.2. Data sources

The Institute for Market Transformation publishes a list of U.S. cities and counties that have adopted BEDB policies (Institute for Market Transformation 2026). I used the January 2026 version of this list to examine all city- and county-level BEDB policies for inclusion in this study. A jurisdiction was included in the study if the following four criteria were met: 1) The text of the local ordinance establishing BEDB was available online; 2) According to the version of the ordinance currently in force, as well as any available prior versions, the square footage threshold determining which commercial buildings are required to comply with the policy has been lowered at least once; 3) The jurisdiction, or the state within which it is located, had not adopted a BPS policy as of 2 years after the lower BEDB threshold came into effect; and 4)

Building-level energy consumption data collected through the BEDB policy is available online for past years in a spreadsheet format.

In total, 11 cities and counties met the above criteria and were included in the study. (Since Montgomery County, Maryland is the only county included in the study, I will refer to all jurisdictions as cities for convenience.) Some cities lowered the threshold for their BEDB policy more than once, so a total of 14 city-year data sets were aggregated for this analysis.

I downloaded all data in the study from websites maintained by city governments. The key variables for the analysis in this paper are gross square footage; total site energy use (TSE), the amount of energy consumed by the building, represented in thousands of British thermal units (kBTU); and building type (for example, "Office," "Non-Refrigerated Warehouse," or "Hotel"). All cities in the study used the U.S. Environmental Protection Agency's ENERGY STAR Portfolio Manager to collect data from building owners. Thus, while the format of the data sets varies slightly between cities, the methodology used to calculate TSE as well as the categories of building types are the same across all cities.

Table 1 lists the cities included in the study. The "previous threshold" and "previous threshold year" columns indicate the prior square footage threshold for compliance with BEDB in each city and the year for which owners of buildings at or above that threshold were first required to report their energy data. The "new threshold" and "new threshold year" columns indicate the lowered threshold for compliance and the first year of energy use data reported under the new threshold. "Data year" represents the year of energy use data that is included in this analysis associated with that threshold-lowering event. The table also reports the number of buildings included in each city-year dataset after the data-cleaning steps described in Section 3.4.

For many cities, "data year" is the same as "new threshold year." In those cities, the data set includes energy use data from buildings between the new threshold and the previous threshold that are in their first year of compliance with BEDB, as well as buildings above the previous threshold that have been subject to BEDB for two or more years. Some cities do not disclose a building's energy use data to the public until the building has been subject to BEDB for two years. For those cities, "data year" is the year after "new threshold year, and the dataset includes buildings between the new and previous thresholds that are in their second year of compliance, as well as buildings above the previous threshold that are in year 3 of compliance or beyond.

Table 1. Cities with BEDB policies, thresholds for compliance, and years of energy data included in this study

City	**Previous threshold (sq. ft.)**	**Previous threshold year**	**New threshold (sq. ft.)**	**New threshold year**	**Data year**	**Number of buildings**
Boston, MA	50,000	2013	35,000	2015	2015	716
Cambridge, MA	50,000	2014	25,000	2015	2015	380
Chicago, IL	250,000	2013	50,000	2014	2015	778
Columbus, OH	100,000	2020	50,000	2021	2022	598
Honolulu, HI	100,000	2022	50,000	2023	2023	161
Honolulu, HI	50,000	2023	25,000	2024	2024	340
Kansas City, MO	100,000	2016	50,000	2017	2018	316
Los Angeles, CA	100,000	2016	50,000	2017	2017	1,358
Los Angeles, CA	50,000	2017	20,000	2018	2018	3,223

Montgomery County, MD	250,000	2015	50,000	2016	2017	302
Portland, OR	50,000	2015	20,000	2016	2017	719
San Francisco, CA	50,000	2010	25,000	2011	2011	597
San Francisco, CA	25,000	2011	10,000	2012	2012	1,042
San Jose, CA	50,000	2018	20,000	2019	2019	501
TOTAL						**11,031**

3.3. Designating treatment and control groups

Because the data used in this study was collected from building owners who were required to report their buildings' energy consumption under the BEDB policy in their jurisdiction, there is no data on buildings that are wholly untreated. In other words, buildings that have never been subject to BEDB are absent from the data set. Instead, the treatment group in this study consists of buildings that have been exposed to treatment for multiple years because they are above the previous threshold for policy compliance, and the control group consists of buildings that are in their first or second of treatment because they are below the previous threshold and above the new threshold that has just gone into effect. While the data does not include any buildings that are entirely untreated, I follow Holland (1986) and others who assert that the difference between a "treated" unit and a "control" unit is that the control unit has been exposed to a different treatment rather than the complete absence of treatment.

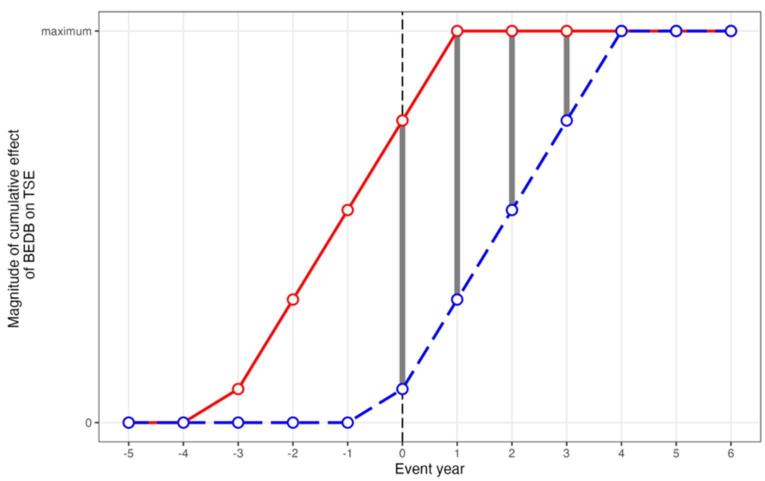

Figure 1. Schematic representation of the effect of BEDB policies on total site energy use (TSE) in buildings in treatment group (red solid line, first treated in year t = -3) and control group (blue dashed line, first treated in year t = 0).

Figure 1 provides a schematic representation of the effect of a BEDB policy over time and its implications for the analysis presented in this paper. The red line reflects buildings in the treatment group that are required to disclose their energy use for the first time in year t = -3. There is a modest treatment effect in the first year, through mechanisms operating on the building owner (e.g., greater attentiveness to energy use). The effect is greater in t = -2, and increases in each of the subsequent three years as consumers respond to the availability of building energy data, expressing a preference for more efficient buildings and incentivizing owners to invest in energy efficiency improvements. At t = 1, buildings in the treatment group have reached the maximum effect of BEDB, and their energy use is unchanged in future years.

Meanwhile, the control group is required to disclose energy use for the first time in year t = 0. Again, the effect is modest in the first year and increases in subsequent years until t = 4, when the maximum effect has been reached and the EUI of treatment and control group buildings is equal. In this paper, I estimated the first gray bar: the difference between treatment and control groups at t = 0. This quantity is less than the cumulative effect of BEDB in the treatment group after three years of exposure to treatment, and it is also less than the maximum effect of BEDB. Nevertheless, an analysis with enough statistical power would confirm the existence and direction of the effect of BEDB on building energy use and place a lower bound on the magnitude of the effect. (For cities where "data year" is the year after "threshold year," I am observing the treatment effect at t=1. While this effect is likely different than the effect at t=0, it should still have a non-zero magnitude as long as the policy has not reached its maximum effect before t=1.)

3.4. Analyzing the pooled data set

To combine each city-year data set into a pooled data set for analysis, I constructed a normalized measure of square footage by subtracted the threshold from the actual floor area for each building. Buildings in the treated group had normalized floor area greater than or equal to 0, while buildings in the control group had normalized floor area less than 0. To control for differences in building energy consumption between cities, I calculated the weighted mean of the natural logarithm of TSE for each city-year dataset around 0 square feet of normalized floor area, using a bandwidth of 20,000 square feet and triangle weighting, and subtracted this value from the observed logarithm of TSE. Thus, in each city-year dataset, buildings near the threshold will have a normalized TSE close to 0.

I excluded from the pooled data set any buildings that were residential or municipally owned, since those buildings were frequently subject to different thresholds and timelines for compliance than commercial buildings. Since some city datasets did not indicate which buildings were municipally owned, I also excluded building types with the highest percentage of municipal buildings, such as K-12 schools, libraries, and police stations. Additionally, I omitted buildings with a square footage below the new threshold (in other words, buildings whose owners voluntarily reported their energy usage), and buildings that reported extreme levels of energy use (greater than 1,000 kBTU / square foot), which likely represented data entry errors.

This paper presents three main analyses. The first model uses the full pooled dataset with no controls. The second model adds controls for the four most common types of buildings in the pooled dataset: office, hotel, non-refrigerated warehouse, and college/university. The third model uses observations only on office buildings, the most common building type. In each analysis, the outcome variable is the normalized natural logarithm of TSE, and the running variable is normalized square footage.

The results from this pooled data set provide a weighted average of the effect of BEDB across multiple geographic contexts and multiple timeframes. BEDB policies vary between cities across several dimensions that plausibly impact the magnitude of the policy's effect on energy use (for example, the ease with which members of the public can access building energy use data and compare across multiple buildings.) While these pooled analyses do not precisely estimate the effect of any particular BEDB policy, a statistically significant finding would confirm that BEDB policies in general have an effect on building energy use.

3.5. Assessing RDD assumptions

To interpret the coefficient from RDD as reflecting the effect of treatment, potential outcomes must be continuous across the treatment threshold. In other words, in the absence of a BEDB policy, buildings on either side of the threshold would have had similar energy consumption; and if buildings on both sides of the threshold had been exposed to BEDB, they would also have similar energy consumption. Additionally, there must be no manipulation of the running variable (square footage) around the cutoff. If building owners were able to manipulate the square footage of their building, they might be able to self-select into the treatment or control group, violating the principle that assignment to treatment should be as-good-as-random.

Figure 2 shows the distribution of buildings in the pooled dataset by normalized floor area in the vicinity of the treatment threshold, in bins of 100 square feet (left) and 10 square feet (right). Each bin includes its left limit but excludes its right limit, so that the bin in the lefthand graph starting at 0 square feet includes buildings with a normalized floor area of 0 square feet but not 100 square feet.

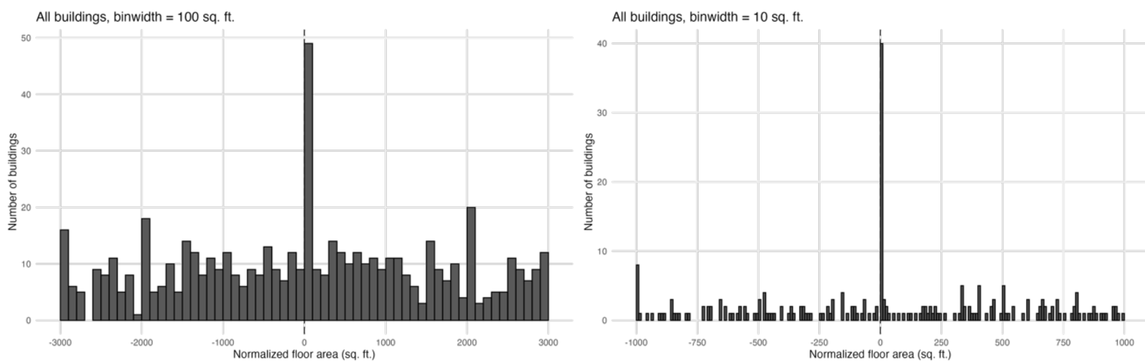

Figure 2. Distribution of buildings by normalized floor area, in bins of 100 square feet (left) and 10 square feet (right). Bins are inclusive of the left limit but not the right limit.

These graphs show a substantial cluster of buildings at or just above 0 normalized square feet. There are 39 buildings in the dataset with a normalized square footage of exactly 0. While this is a small share of the 11,000-plus observations in the dataset, a cluster of observations at or near the threshold could bias the estimate of the treatment effect, particularly if this clustering results from self-selection into treatment or indicates a discontinuity in potential outcomes.

There are several reasons to think that the cluster of observations at normalized square footage = 0 does not reflect self-selection into treatment. First, to the extent that building owners prefer not to be subject to BEDB policies because of the time and effort associated with reporting energy consumption, we would expect owners to misreport their buildings' square footage in order to fall below the threshold, rather than precisely at the threshold. Buildings with a square footage exactly equal to the threshold are generally required to comply with BEDB policies. Second, the floor area information used to determine whether a building is required to comply with BEDB requirements is generally sourced from a municipal database that preexists the adoption of BEDB (for example, the city assessor's database). Third, there are clusters of buildings around other round values of normalized square footage outside of the range represented in Figure 2. When the dataset is divided into bins of width = 10 square feet, as in the righthand graph in Figure 2, bins that include normalized square footage values divisible by 10,000 (excluding the bin with square footage = 0) have a mean of 31.7 observations, and bins that include square footage values divisible by 5,000 but not divisible by 10,000 have a mean of 18.6 observations. These are substantially higher than the mean of 1.4 observations in bins that do not include values divisible by 100.

Under BEDB policies, building owners realize no advantage from having a normalized square footage value that is exactly divisible by 5,000 or 10,000. Thus, the higher frequency of observations in these bins implies that there is a tendency to assign buildings a square footage value that is a round number, rather than a deliberate effort to avoid the requirements of BEDB policies. Nevertheless, to address the possibility that these clusters introduce discontinuities in potential outcomes, supplemental analyses in Section 4.2 omit buildings with normalized square footage exactly equal to 0 or divisible by 5,000 or 1,000.

I also investigated whether the proportion of buildings of the four most common types is similar on both sides of the threshold. A discontinuity in a covariate such as building type could indicate that potential outcomes are not continuous across the threshold, leading to bias in the estimate of the treatment effect. Using a regression discontinuity model to estimate changes in the frequency of building types across the treatment threshold, I found a decrease of approximately 5% in the share of office buildings (p-value: .103) and 3% in the share of hotel buildings (p-value: .059) among buildings just above the threshold, compared to buildings below the threshold. Estimates of the change in frequency of warehouses and college/university buildings were below 0.5% and not statistically significant. To account for potential bias introduced by discontinuities in the distribution of building types across the threshold, model 2 includes controls for common building types, and model 3 is restricted to office buildings.

4. Results

4.1. Main analyses

Table 2 presents the results of three RDD models estimating the effect of BEDB on building energy use. For each model, the coefficient is the estimated treatment effect, the change in the natural logarithm of TSE at the threshold. Negative values of the coefficient mean that energy use is lower in the treatment group than the control group. Model 1 uses all observations in the pooled dataset with no controls, while Model 2 adds controls for the four most common types of buildings: office, hotel, non-refrigerated warehouse, and college/university. Model 3 is restricted to office buildings. All coefficients are bias-corrected, and confidence intervals and p-values are determined using the robust methodology in *rdrobust*.

Figure 3 displays the results of Model 1 (left) and Model 3 (right) using the conventional method for estimating RDD models (without the *rdrobust* bias correction methodology), along with the mean of the outcome variable evaluated in bins of 500 square feet. The vertical gap between the two regression lines at normalized floor area = 0 represents the estimated treatment effect under the conventional method. These conventional values differ from the values presented in Table 2, which uses a robust estimation method that reduces finite-sample bias. The estimates in Table 2 are likely more accurate, and Figure 3 is presented for illustrative purposes only.

Table 2: Results for three regression discontinuity models assessing the effect of BEDB policies on the natural logarithm of TSE

Model	1	2	3
Coefficient (95% confidence interval)	-0.234 (-0.407, -0.061)	-0.216 (-0.383, -0.049)	-0.141 (-0.347, 0.065)
P-value	0.008	0.011	0.180
Bandwidth	19,957	20,517	28,933
N within bandwidth	3,475	3,626	1,515
All building types	x	x	
Office buildings only			x
Controls for common building types		x	

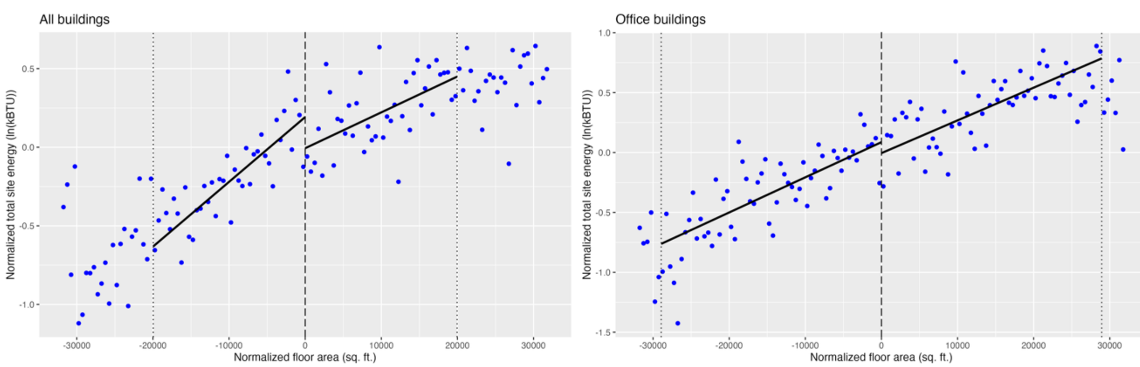

Figure 3. Regression discontinuity results for Model 1 (left) and Model 3 (right). Solid black lines represent the regression model. The vertical dashed line is the treatment threshold, and the vertical dotted lines are located at the upper and lower limits of the bandwidth used to estimate the model. Blue dots are the mean of the outcome variable calculated in bins of 500 square feet.

In Model 1, the estimated treatment effect is 0.234 log points, corresponding to a 20.9% decrease in TSE in the treatment group compared to the control group in the vicinity of the threshold. Model 2, which includes building type controls, estimates an effect of similar magnitude: 0.216 log points or a 19.4% decrease in TSE. When the dataset is restricted to office buildings in Model 3, the estimated effect is reduced to 0.141 log points or a 13.2% decrease in TSE. The results from Model 1 and Model 2 are statistically significant at the $p = 0.05$ level, while the result from Model 3, with a smaller effective sample size, is not significant.

4.2. Sensitivity analyses

Table 3 presents the results from Model 1 alongside several variant models intended to evaluate the robustness of the findings from the main analyses and address potential sources of bias. Models 1a – 1c test whether omitting clusters of observations at round values of square footage changes the estimated treatment effect. Model 1a omits buildings with normalized square footage of exactly 0, and Models 1b and 1c omit buildings with normalized square footage divisible by 5,000 or by 1,000, respectively.

While the bandwidth in Model 1 is selected according to the *rdrobust* method to minimize error in estimating the treatment effect, the chosen bandwidth (19,957) is wide relative to the range of the data. Models 1d – 1f use bandwidths of 15,000, 10,000, and 5,000 square feet, respectively, to evaluate how using a narrower bandwidth changes the estimated treatment effect.

Table 3: Results for Model 1 and six variant regression discontinuity models assessing the effect of BEDB policies on the natural logarithm of TSE

Model	**1**	**1a**	**1b**	**1c**	**1d**	**1e**	**1f**
Coefficient (95% confidence interval)	-0.234 (-0.407, -0.061)	-0.188 (-0.362, -0.014)	-0.174 (-0.347, 0.000)	-0.193 (-0.372, -0.014)	-0.259 (-0.507, -0.011)	-0.226 (-0.521, 0.070)	-0.018 (-0.420, 0.384)
P-value	0.008	0.034	0.050	0.035	0.041	0.135	0.930
Bandwidth	19,957	20,508	21,086	21,616	15,000	10,000	5,000
N within bandwidth	3,475	3,586	3,452	3,247	2,683	1,732	893
All building types	x	x	x	x	x	x	x
Controls for common building types							
Sq. ft. = 0 omitted		x	x	x			
Sq. ft. divisible by 5,000 omitted			x	x			
Sq. ft. divisible by 1,000 omitted				x			
Bandwidth restricted					x	x	x

Omitting observations with round values of square footage does not substantially change the results. The estimates from Models 1a – 1c are statistically significant at the p = 0.05 level and similar in direction and magnitude to model 1. Model 1d, with a bandwidth of 15,000, also returns a similar, statistically significant estimate of the treatment effect. However, Model 1f, with the narrowest bandwidth, returned a near-zero point estimate of the treatment effect with a P-value of 0.930. These results suggest that the estimated treatment effect is unstable for narrow bandwidths. This is not necessarily an indication that the treatment effects estimated in the main models are incorrect; rather, a high variance in building energy use at any given square footage value, combined with a small effective sample size, may make it difficult to detect a treatment effect when the bandwidth is restricted.

5. Discussion and conclusion

Using RDD on a pooled data set including building-level energy data from 11 U.S. cities that have adopted BEDB policies, I found a statistically significant difference in energy consumption between a treatment group of buildings that have been subject to these policies for multiple years and a control group of buildings that are in their first or second year of compliance. Across all types of commercial buildings, buildings in the treatment group used

19.4–20.9% less energy compared to buildings in the control group. The estimated reduction in energy use was 13.2% when the data set was restricted to office buildings only, and this result was not statistically significant due to a smaller sample size.

These results must be viewed as preliminary. Previous studies using plausible causal methodology have generally found much smaller effects of BEDB on energy use. For example, the preferred specification in Palmer and Walls (2015b) found that BEDB reduced utility bills by 3%.

For RDD to provide an unbiased estimate of the treatment effect, observations on either side of the treatment threshold must be similar along all dimensions affecting the outcome except for their assignment to the treatment or control group. While some of the models in this study control for one potential confounding variable, building type as defined in the U.S. Environmental Protection Agency's ENERGY STAR Portfolio Manager, they do not account for all factors that may differ between buildings in the treatment and control groups. For example, zoning ordinances may limit the size of buildings that may be used for certain purposes. If the zoning code in one city allows for 24-hour grocery stores to be built up to 30,000 square feet, we might expect to see a concentration of such stores just below that size. If 30,000 square feet is also the threshold for compliance with the city's BEDB policy, the existence of 24-hour grocery stores on only one side of the threshold could introduce bias into the estimated treatment effect. The building types included in this study do not necessarily capture such fine-grained differences in how buildings are used.

Additionally, cities may adopt other policies that use the same threshold as BEDB and also affect building energy use. This study accounts for one policy, BPS, that is commonly adopted alongside BEDB, by omitting any cities from the pooled dataset that adopted BPS at the same time as BEDB or within 2 years after a change in the BEDB threshold came into effect. However, other policies — such as incentives for energy efficiency upgrades — may use the same threshold as BEDB, and this study does yet take those policies into account.

Although the results in this paper are not definitive, the study includes several key methodological innovations that will contribute to future research on BEDB and other information disclosure policies. First, I aggregated data from nearly a dozen cities to increase statistical power. Subtracting the square footage value of the threshold from the reported floor area of each building allows for data from cities with different BEDB thresholds to be pooled together for the purpose of RDD analysis.

Second, I overcame the core problem in studying BEDB with the data generated by the policies — namely, a lack of data on untreated buildings — by using data from cities where compliance thresholds that have been lowered over time. This approach allows for comparisons between buildings that have been exposed to a greater versus lesser duration of treatment. While it does not estimate the magnitude of the effect of BEDB on treated buildings compared to wholly untreated buildings, it does estimate a lower bound on the effect size, assuming that the treatment effect increases over time during the first few years the policy is in effect.

Finally, this study used RDD to increase confidence in the causal interpretation of the results, compared to prior studies conducted with difference-in-difference designs. With additional work to account for potential sources of bias, this approach can generate robust estimates of the effect of BEDB on energy consumption. These findings will empower public officials, nonprofit leaders, and industry stakeholders to design and implement information disclosure policies with greater confidence that these policies will promote the desired end goal: a reduction in building energy consumption.

References

Allcott, Hunt, and Michael Greenstone. 2012. "Is There an Energy Efficiency Gap?" *Journal of Economic Perspectives* 26(1): 3–28. doi:10.1257/jep.26.1.3.

Ayres, Ian, Sophie Raseman, and Alice Shih. 2009. *Evidence from Two Large Field Experiments That Peer Comparison Feedback Can Reduce Residential Energy Usage*. Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w15386.

Beddingfield, Erin, Julie Hughes, and Zachary Hart. 2017. *Putting Data to Work: How Cities Are Using Building Energy Data to Drive Efficiency*. Institute for Market Transformation. https://imt.org/resources/putting-data-to-work-how-cities-are-using-building-energy-data-to-drive-eff/.

Brounen, Dirk, and Nils Kok. 2011. "On the Economics of Energy Labels in the Housing Market." *Journal of Environmental Economics and Management* 62(2): 166–79. doi:10.1016/j.jeem.2010.11.006.

Calonico, Sebastian, Matias D Cattaneo, and Max H Farrell. 2020. "Optimal Bandwidth Choice for Robust Bias-Corrected Inference in Regression Discontinuity Designs." *The Econometrics Journal* 23(2): 192–210. doi:10.1093/ectj/utz022.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik. 2015. "Rdrobust: An R Package for Robust Nonparametric Inference in Regression-Discontinuity Designs." *The R Journal* 7(1): 38–51.

Costa, Dora, and Matthew Kahn. 2010. *Energy Conservation "Nudges" and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment*. Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w15939.

Department of the Environment, Water, Heritage and the Arts. 2008. *Energy Efficiency Rating and House Price in the ACT*. Australian Government, Department of the Environment, Water, Heritage and the Arts. https://www.exemplary.com.au/download/eer-house-price-act.pdf.

Fuerst, Franz, and Patrick McAllister. 2011. "Green Noise or Green Value? Measuring the Effects of Environmental Certification on Office Values." *Real Estate Economics* 39(1): 45–69. doi:10.1111/j.1540-6229.2010.00286.x.

Gilmer, Robert W. 1989. "Energy Labels and Economic Search." *Energy Economics* 11(3): 213–18. doi:10.1016/0140-9883(89)90026-1.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396): 945–60. doi:10.1080/01621459.1986.10478354.

Institute for Market Transformation. 2026. "Comparison of U.S. Commercial Building Energy Benchmarking and Transparency Policies." https://imt.org/resources/comparison-of-commercial-building-benchmarking-policies/ (March 13, 2026).

Krukowski, Andrea, and Andrew C. Burr. 2012. *Energy Transparency in the Multifamily Housing Sector: Assessing Energy Benchmarking and Disclosure Policies*. Institute for Market Transformation. https://imt.org/resources/energy-transparency-in-the-multifamily-housing-sector/.

Laquatra, Joseph. 1986. “Housing Market Capitalization of Thermal Integrity.” *Energy Economics* 8(3): 134–38. doi:10.1016/0140-9883(86)90011-3.

Meng, Ting, David Hsu, and Albert Han. 2017. “Estimating Energy Savings from Benchmarking Policies in New York City.” *Energy* 133: 415–23. doi:10.1016/j.energy.2017.05.148.

Miller, Norm, Jay Spivey, and Andrew Florance. 2008. “Does Green Pay Off?” *Journal of Real Estate Portfolio Management* 14(4): 385–400. doi:10.1080/10835547.2008.12089822.

Mims, Natalie, Steven R Schiller, Elizabeth Stuart, Lisa C Schwartz, Chris Kramer, and Richard Faesy. 2017. *Evaluation of U.S. Building Energy Benchmarking and Transparency Programs: Attributes, Impacts, and Best Practices*. Lawrence Berkeley National Laboratory. https://emp.lbl.gov/publications/evaluation-us-building-energy.

Nadel, Steven, and Adam Hinge. 2020. *Mandatory Building Performance Standards: A Key Policy for Achieving Climate Goals*. ACEEE. https://www.aceee.org/white-paper/2020/06/mandatory-building-performance-standards-key-policy-achieving-climate-goals.

Palmer, Karen L., and Margaret Walls. 2015a. “Can Benchmarking and Disclosure Laws Provide Incentives for Energy Efficiency Improvements in Buildings?” *SSRN Electronic Journal*. doi:10.2139/ssrn.2564251.

Palmer, Karen L., and Margaret Walls. 2015b. “Does Information Provision Shrink the Energy Efficiency Gap? A Cross-City Comparison of Commercial Building Benchmarking and Disclosure Laws.” *SSRN Electronic Journal*. doi:10.2139/ssrn.2622692.

Stavins, Robert N. 2003. “Experience with Market-Based Environmental Policy Instruments.” In *Handbook of Environmental Economics*, Elsevier, 355–435. doi:10.1016/S1574-0099(03)01014-3.

Weil, David, Archon Fung, Mary Graham, and Elena Fagotto. 2006. “The Effectiveness of Regulatory Disclosure Policies.” *Journal of Policy Analysis and Management* 25(1): 155–81. doi:10.1002/pam.20160.

Woodson, Colleen, Jon Braman, Janne Flisrand, William Weber, Patrick Smith, and Katelyn Tsukada. 2015. *EnergyScoreCards Minnesota: Results from Energy and Water Benchmarking in 500+ Minnesota Multifamily Buildings*. Bright Power, Inc. https://web.archive.org/web/20170224021705/http://mn.gov/commerce-stat/pdfs/bright2016-mfbenchmark-final.pdf.