

Utilities Included: Split Incentives in Commercial Electricity Contracts

Katrina Jessoe,^a Maya Papineau,^b and David Rapson^c

ABSTRACT

This paper quantifies a tenant-side “split incentives” problem that exists when the largest commercial sector customers are on electricity-included property lease contracts, causing them to face a marginal electricity price of zero. We use exogenous variation in weather shocks to show that the largest firms on tenant-paid contracts use up to 14 percent less electricity in response to summer temperature fluctuations. The result is retrieved under weaker identifying assumptions than previous split incentives papers, and is robust when exposed to several opportunities to fail. The electricity reduction in response to temperature increases is likely to be a lower bound when generalized nationwide and suggests that policymakers should consider a sub-metering policy to expose the largest commercial tenants to the prevailing retail electricity price.

Keywords: Principal-Agent Problem, Split Incentive, Contracts, Commercial Sector, Electricity

<https://doi.org/10.5547/01956574.41.5.kjes>

1. INTRODUCTION

Separating the party who pays for energy from the one making decisions about electricity use has long been cited as creating incentives for energy over-consumption or underinvestment in energy efficiency in both the commercial and residential sectors. In the U.S., roughly 20 percent of commercial building occupants rent space with electricity bundled into their monthly rent. Under this contract structure commercial tenants face zero marginal cost of consuming electricity, creating an incentive to over-consume. The remaining 80 percent of tenants pay their own monthly utility bills, which will dampen the incentive for building owners to invest in energy efficiency if owners cannot capitalize on a rent premium for energy efficiency upgrades. These misalignments between tenant and landlord incentives may lead to overconsumption of energy and overproduction of pollution that Pigouvian taxes are not well suited to correct (Jaffe and Stavins, 1994; Gillingham and Palmer, 2014). Given that the commercial sector accounts for over 35 percent of end-use electricity consumption in the U.S., the welfare costs from excess energy use may be substantial. Yet little evidence exists about the magnitude of these “split incentive” principal-agent problems in the commercial sector.¹

1. More broadly, the existence of this principal-agent problem may justify programs or regulations to mandate energy efficiency, such as building energy standards or the use of firm-level energy saving obligations, also known as “white certificates”, that have been adopted in several European countries (Stavins, 2011; Giraudet and Finon, 2015; Papineau, 2017).

a Department of Agricultural and Resource Economics, UC Davis. E-mail: kkjessoe@ucdavis.edu.

b Department of Economics, Carleton University. E-mail: maya.papineau@carleton.ca.

c Corresponding author. Department of Economics, UC Davis. E-mail: dsrapson@ucdavis.edu.

In this paper, we estimate an important component of the change in electricity use from switching commercial customers on electricity-inclusive rent contracts to tenant-paid utility contracts, a distinction we refer to as “contract type”. We do this by evaluating how the relationship between electricity use and temperature (the temperature response gradient, henceforth “TRG”) differs by contract type. We illustrate how the structure of the rental contract may create two distinct split incentives, one on the intensive and another on the extensive margin of demand for energy services, which lead to different empirical predictions relating to the TRG. When considering the intensive margin, the TRG will be less steep under a tenant-paid than an electricity-inclusive, or “owner-paid” contract. This occurs because, for a given level of energy efficiency capital, firms on a tenant-paid contract pay a positive marginal price for electricity use while those on an owner-paid contract face a marginal price of zero. The second split incentive relates to owner incentives to invest in energy efficient durables. Under a tenant-paid contract, owners have little incentive to invest. Since (all else equal) lower investment in energy efficiency leads to a steeper TRG, firms located in buildings on tenant-paid contracts should exhibit a steeper TRG relative to owner-paid contracts. We refer to this as the extensive margin effect. These two split incentives impact the TRG in opposite directions in relation to contract type, allowing us to empirically test which split incentive (if any) dominates in our setting.

Results suggest that the intensive margin effect dominates among the largest firms. Tenant-paid contracts induce considerable energy savings among these customers during the hottest summer months. For the largest decile of firms, switching from an owner-paid to tenant-paid utility contract would reduce electricity use by roughly 3 percent over the course of a year and up to 14 percent in the summer months. The annual savings among large consumers are comparable to popular energy conservation measures such as home energy reports, which produce average savings of approximately 2 percent (Allcott, 2011). Furthermore, the savings occur at times when the value of electricity is likely to be high: during the hottest days of the year. Our finding that the largest customers are most responsive to contract type corroborates recent evidence from the residential sector in Sweden (Elinder et al., 2017). In contrast, contract type does not measurably impact consumption decisions for the smallest 90 percent of commercial customers. These results are consistent with profit-maximizing firms facing adjustment costs in electricity consumption, such that a relatively small absolute value of bill savings, among smaller firms, would not warrant conservation behavior.

Our empirical approach exploits the differential effect of an exogenous weather shock on electricity use across firms on an owner versus tenant-paid contract. To do this we make use of cross-sectional variation in local weather exposure within a calendar billing month generated from the staggering of electricity billing periods across customers. We combine these weather data with monthly bills from 1,074 commercial firms serviced by a Connecticut electric utility between October 2007 and May 2011, and property-level information on fixed observables including whether the tenant or landlord pays the electric bill. This panel data set allows us to examine the differential impact of local weather shocks on electricity use across contract types, controlling for potential selection into contract type based on firm or fixed building attributes. A “levels” comparison of electricity use across contract type would be biased if firms on owner- and tenant-paid contracts differ in ways that are correlated with energy use. Focusing instead on the TRG across contract type captures the sensitivity of electricity consumption to fluctuations in temperature and allows us to control for a rich set of variables that may correlate with selection into contract type, thereby permitting identification under weaker assumptions.

Our identifying assumption is that selection into contract type is unrelated to unobservable electricity demand drivers that are correlated with the TRG. We present three pieces of empirical evidence that support this assumption. First, motivated by recent work demonstrating that the elec-

tricity response to temperature shocks meaningfully differs across certain building attributes, we directly control for the possibility that the TRG is heterogeneous in observable building attributes (Novan et al., 2017). After controlling for interactions between temperature and attributes such as building age and industry type, our results are unchanged. Second, we use a change to a Connecticut metering regulation. This change was legislated after the end of our sample period and altered building owners' ability to select into contract type. It provides us with an opportunity to examine the TRG of firms located in buildings that switched contract types shortly after the change, and to test whether they exhibit a differential response gradient. They do not. Third, we assess the effect of potential correlations between any remaining unobservable characteristics and the treatment, as described in Oster (2017). This places bounds on the potential bias from selection on unobservables. Each of these tests exposes our identifying assumption to an opportunity to fail, and the results of each test support our main conclusions.

Given the size of the firms responsive to contract type, the estimated treatment effect translates into significant costs from misaligned incentives. Using very conservative assumptions, we find that if incentives were aligned among the largest decile of commercial customers nationwide, total energy savings would be roughly one and a quarter times the savings from solving the split incentives problem for the entire U.S. residential electricity sector. The magnitude of the treatment effect and the relative size of large commercial firms are the primary factors leading to this result. Though the number of commercial customers affected by the split incentives problem is small relative to residences, these customers use much more energy. Thus, addressing the commercial split incentive problem requires a fraction of the contact points, while likely leading to greater energy savings. Our estimates imply greenhouse gas reductions of between 615-1200 thousand tons of CO₂ per year, or (to give a sense of scale) roughly 3.3 to 6.6 times the average annual savings from yearly Weatherization Assistance Program retrofits. These savings may be achievable at a relatively low cost. When we compare the cost of retrofitting units with sub-meters (to allow switching to tenant-paid utility bills) with the estimated annual bill savings amongst the largest customers, the payback period is less than one year. One caveat to note is that our results are identified based on the TRG. A switch from owner to tenant-paid contracts, and the accompanying disincentive for landlords to invest in energy efficiency, may also alter the level of electricity use. While many building-level energy efficiency investment choices available to landlords affect the TRG, we cannot rule out that these "level" effects may lead to a consumption increase and mitigate or even overwhelm the TRG savings we identify relating to the temperature gradient.

Despite the robustness of our results, identification of our main effect arises from the behavior of 110 large firms, 19 of whom are on owner-paid contracts. Our need to lean on a small sample raises the possibility that our study may be under-powered and that our results may be driven by outliers. While our research setting prevents us from completely nullifying these concerns, we present evidence that allows us to reject the hypothesis that our findings are driven by a single outlier among the largest owner-paid firms in our sample.

This work makes four main contributions to the academic literature and environmental policy discussion. First, compared to the residential setting where a growing literature points to both the potential and limitations of energy efficiency and contracting solutions (Gillingham et al., 2012; Hassett and Metcalf, 1999; Fowlie et al., 2015; Elinder et al., 2017), little is known about the commercial setting. We provide a commercial counterpart to existing residential estimates on the split incentives problem. Second, our identification strategy makes several advances towards credibly estimating the magnitude of the split incentives problem. The response gradient, temperature-characteristic interactions, contract switcher controls, and Oster bounds each provide support for the identifying assumption and extend the existing literature on split incentives. Third, our re-

sults reveal substantial heterogeneity in firm responsiveness to contract type and point to the importance of looking beyond population average treatment effects. Lastly, our results suggest a targeted prescriptive policy of tenant-paid contracts among large firms may be a net beneficial greenhouse gas abatement strategy.

The rest of the paper is organized as follows. Section 2 presents our conceptual framework and our predictions about the TRG among tenant- and owner-paid utilities under the two potential split incentives. Section 3 reviews our empirical setting and presents our data. Section 4 discusses identification and our empirical specifications. Section 5 presents our empirical results and policy implications. Section 6 briefly concludes.

2. CONCEPTUAL FRAMEWORK

The two forms of the split-incentive principal-agent problem introduced in the previous section translate into simple empirical predictions, which we articulate in more detail in this section. The structure of a rental contract shifts both the responsibility for electricity bill payments (the intensive margin) and the extent of investment in energy efficiency (the extensive margin). These, in turn, affect how we would expect electricity consumption to change in response to an exogenous demand shifter such as temperature fluctuations. Our framework highlights that when the intensive margin dominates, tenant-paid rental contracts will lead to electricity consumption levels that exhibit a weaker positive covariance with temperature relative to owner-paid contracts. In contrast, if the extensive margin dominates, the opposite holds: owner-paid contracts create incentives that will result in a lower consumption response to temperature changes. Our setting allows for a test of which of these channels dominates.

2.1 Set-up and Notation

Tenant k 's electricity consumption, measured in kilowatt-hours (kWh), in a building owned by landlord j is given by $Y = f(E, T, P_k(W))$. Electricity use depends on energy efficiency capital, E , temperature, T , and the price per kWh of electricity paid by tenant k , $P_k(W)$, which is a function of contract type W . Under standard conditions, electricity consumption is decreasing in P and E , and increasing in T .² For any building there are two potential contract types offered: an owner-paid contract, $W = 0$, and a tenant-paid contract $W = 1$. Under the former $P_k(W) = P_k(0) = 0$, and under the latter $P_k(W) = P_k(1) = p$, where p is the retail price of electricity. When choosing where to rent commercial space, tenants choose the contract type and the capital characteristics of a building. Conditional on these decisions, which are assumed to be fixed in the short-run, tenants make monthly decisions about electricity use.

2.2 Competing Split Incentive Hypotheses

Our empirical predictions relate to the temperature response gradient (TRG): $\frac{\partial Y}{\partial T}$. The TRG reflects the sensitivity of electricity consumption to fluctuations in temperature, and it has two attractive features. First, since temperature fluctuations are exogenously experienced by tenants, this variation is an appealing source for identification. Second, as we will discuss, the TRG will differ as a function of contract type, W , along two margins: an extensive and an intensive one.

2. These follow, respectively, from a downward-sloping demand curve, an assumption that rebound effects are less than 100 percent in magnitude, and by restricting our attention to temperatures of 65F and above such that derived electricity demand (via air conditioning) is increasing in temperature.

Temperature interacts with energy efficiency and price to influence electricity use. The more energy efficient a building is, the less sensitive energy demand is to temperature increases (i.e. the TRG is shallower): $\frac{\partial^2 Y}{\partial T \partial E} < 0$. This intuition is supported by recent empirical work demonstrating that the electricity response to temperature shocks is decreasing in the stringency of building energy codes (Novan et al., 2017). The TRG will also be dampened at higher prices, all else equal. This follows from the costs associated with increasing air conditioning as temperature increases and as the tenant pays a higher price for electricity: $\frac{\partial^2 Y}{\partial T \partial P_k} < 0$. The incentives created by contract type flow from these pieces of intuition.

When $W = 1$, tenants pay $p > 0$ for every kWh of electricity they use. In contrast, under an owner-paid contract, the marginal price tenants pay for electricity equals 0. Because of this difference in prices, tenant k 's TRG will be shallower under a tenant-paid than under an owner-paid contract, all else equal (including E). Our first empirical prediction follows.

Prediction 1:

$$\left. \frac{\partial Y}{\partial T} \right|_{E, p > 0} < \left. \frac{\partial Y}{\partial T} \right|_{E, p = 0} \quad (1)$$

The level of E may also be expected to change with contract type, and this will have a distinct effect on the TRG. Consider landlord j 's decision to invest in energy efficiency capital. Landlord j chooses E to maximize profits from the rental of commercial units,

$$\max_E \pi_j = R(E) - P_j(W)Y(E) - rE.$$

Rental revenues denoted by R depend on energy efficiency capital E , where we assume that rents are weakly increasing in E . Total costs comprise the electricity bill, $P_j(W)Y(E)$, which is a function of the price per kWh of electricity paid by landlord j and the quantity of electricity consumed by the tenant. Total costs also include the costs to purchase energy efficiency capital, where we assume the cost per unit of capital is r . Under differentiability, the landlord chooses a quantity of E such that

$$\frac{\partial R}{\partial E} - P_j(W) \frac{\partial Y}{\partial E} = r. \quad (2)$$

The landlord chooses to invest in energy efficiency capital up to the point where the increase in marginal benefit from a unit of energy efficiency capital is equal to the marginal cost. The marginal benefit from energy efficiency capital consists of two components: the increase in rents from additional energy efficiency capital and the reduction in the electricity bill from a decrease in electricity consumption.

Now suppose that a building with a landlord-paid contract structure is exogenously placed on a tenant-paid contract, $W = 1$. The tenants occupying this building will now pay a price p per kWh of electricity used, and the price paid by the landlord is $P_j(W) = P(1) = 0$. As shown below, this reduces the incentive for a landlord to invest in energy efficiency because the landlord will no longer benefit from lower electricity bills after investing in energy efficiency, and equation (2) will simplify to

$$\frac{\partial R}{\partial E} = r. \quad (3)$$

If the landlord cannot obtain a rent premium that fully accounts for the electricity bill reductions from energy efficiency investments, her demand for energy efficiency capital will be reduced. This leads to the second split incentives problem: if tenant k is on a tenant-paid contract, the building owner will be disincentivized from investing in energy efficiency capital.

This outcome leads to our second prediction about the TRG and how it changes with contract type. Let L denote a low level of E , and H denote a high level.

Prediction 2:

$$\left. \frac{\partial Y}{\partial T} \right|_{E=L, p} > \left. \frac{\partial Y}{\partial T} \right|_{E=H, p} \quad (4)$$

If the split incentive effect is operating entirely through the extensive margin (i.e. holding electricity price constant), then tenant k 's TRG will be steeper if $W = 1$ than if $W = 0$. This prediction arises because of lower levels of investment in energy efficiency capital by landlord j in a tenant-paid building.

The two split incentive effects set forth in predictions 1 and 2 represent competing hypotheses about the TRG in tenant-paid contract buildings. These two effects are likely to occur simultaneously, such that any conservation benefit from a tenant paying their own energy bills could be mitigated or overwhelmed by building capital inefficiencies. Alternatively, in owner-paid buildings attenuation in the temperature gradient from investment in energy efficiency capital may be compromised or dominated by split incentives from the absence of a price signal for tenants. What we observe in our data is the net outcome of these two competing effects: a negative estimate for the impact of a tenant-paid contract on electricity consumption indicates that the intensive margin dominates, whereas a positive coefficient indicates that the extensive margin dominates.

3. EMPIRICAL SETTING: BACKGROUND AND DATA

The split incentives literature goes back several decades and has used a number of different methodological approaches, from interviews with building industry professionals (Blumstein et al., 1980), to engineering approaches combined with survey data (Murtishaw and Sathaye, 2006) and regression-based analyses (Levinson and Niemann, 2004; Davis, 2012; Gillingham et al., 2012; Krishnamurthy and Krström, 2015; Elinder et al., 2017). Most recently, energy savings resulting from a switch from owner- to tenant-paid contracting in the residential sector have been documented in a quasi-experimental setting by Elinder et al. (2017). The split incentive problem resulting from tenant-paid electricity contracts, which reduce the incentive for landlord investment in energy efficiency capital, has been documented by Davis (2012), Krishnamurthy and Krström (2015) and Myers (2015).

In the residential sector, existing studies have found that the split incentive effect on aggregate consumption is of modest to moderate magnitude. Levinson and Niemann (2004) find that energy bills in the U.S. are 0.7 percent higher when apartment dwellers do not pay for heat. Elinder et al. (2017) find that electric energy consumption in a sample of Swedish apartments falls by close to 25 percent when tenants are shifted from owner-paid to tenant-paid electricity billing, where this effect is driven in large part by the highest-consuming households.³

3. Another dimension to the principal-agent problem is less than efficient turnover from oil-fired to gas-fired boilers for residential heating in the northeastern U.S. (Myers, 2015). This outcome is consistent with asymmetric information over heating costs when tenants pay for heat. Inefficient turnover led to 37 percent higher annual heating costs in the 1990–2009 period.

While the engineering literature has identified several channels through which split incentives may affect commercial sector consumption, a gap remains in our understanding of its precise magnitude. One exception is Kahn et al. (2014), who find that energy consumption by tenants who pay their own energy bills is 20 percent lower compared to owner-paid units. However, as noted by the authors, this estimate reflects the effect of both contract type itself, and selection into contract type and buildings based on preferences for energy services.

In this study we focus on energy overconsumption that arises from air conditioning. The mechanism of control over air-conditioning by tenants is an important consideration in measuring firm-level response to contract type. Individual tenants in multi-tenant buildings may each exercise control over cooling in their units through zonal air-conditioning, which has been commercially available since the late 1950s (Meyer, 2006; Hoyer, 2014; AHRI, 2017). In the commercial sector in the summer months, buildings can be over-cooled, leading to reduced comfort levels and an accompanying increase in electricity consumption of up to 8 percent (Derrible and Reeder, 2015). Equipment and electronics usage may also increase if there are poor incentives to conserve. Sanchez et al. (2007) find that office equipment and electronics—such as computers, personal space heaters and fans—account for up to 20 percent of annual building-level electricity consumption.

We evaluate our research questions within the jurisdiction of United Illuminating (UI), an investor-owned electric utility in Connecticut servicing customers across 17 counties. Figure 1 shows its service territory. The regulations surrounding metering in Connecticut make it an advantageous setting in which to study the split incentives problem. To get a sense for the regulatory

Figure 1: UI Territory



Notes: United Illuminating's service territory. It offers electricity distribution services to 17 counties in Connecticut, an area totalling 335 square miles.

landscape, consider the owner of a multi-tenanted building. Monitoring each tenant's individual electricity use would require the installation of a sub-meter. However, prior to the summer of 2013 the state prohibited the retrofitting of commercial and multi-family buildings with sub-meters. As a result, only buildings initially constructed with sub-meters in place could charge individual tenants for energy consumption.⁴ In all other buildings electricity consumption was master-metered (monitored at the building level), and thus tenants signed owner-paid contracts. Since our analysis focuses on the time period 2007 to 2011, the presence of sub-meters in buildings is predetermined from the perspective of current owners and tenants. While tenants are able to choose buildings based on electricity contract type, doing so limited their choice set to buildings retrofitted with a sub-meter at the time of construction.

In 2013, new legislation passed by the Connecticut General Assembly eliminated the sub-metering prohibition (Hartford Business Journal, 2013). While we cannot directly test the effect of this change on electricity use due to the fact that it post-dates our electricity billing sample, the legislative change enables us to gain further insights into selection on contract type based on firm and building-level energy consumption characteristics.⁵

3.1 Data

We combine three data sets to form a panel of 39,233 observations on monthly electricity use from 1,074 firms, and monthly weather from 32 unique zip codes.⁶ The first data source is monthly billing data provided by UI that reports account-level monthly electricity consumption (in kWh), peak monthly throughput or power (in kW), and monthly expenditure. These data also contain information on the industrial classification number—or NAICS code—of each account. The second source is the CoStar Group, a commercial-sector multiple listing service and database that includes building-level information on utility contracts and hedonic characteristics, such as year of construction, number of stories and total square feet. Third, we obtained average daily temperature data from the National Oceanic and Atmospheric Administration (NOAA).

Table 1 presents sample summary statistics on monthly consumption (in kWh), peak throughput in the highest demand hour within a billing month (kW), location, and industry, by contract type, for the firms in our sample. The average customer (across contract types) spends about \$695 a month on electricity; the average building is approximately three stories; and the primary industry is 'Finance, Real Estate and Management', which makes up about 50 percent of the sample among both contract types.⁷ The predominant share of accounts are located in office buildings (72 percent), followed by industrial buildings (22 percent). In our sample, about 84 percent of firms pay their own electricity bill. The regional distributions are also similar across contract types, with about 30 percent of observations in central cities, and the rest located in more suburban areas.

Our empirical analysis will seek to test whether the effect of contract type on electricity use differs by firm size, given that absolute bill reductions from conservation will be concentrated among the largest firms. To measure firm size, we follow a similar approach to Aigner and

4. Several states have historically banned utility sub-metering, primarily for consumer protection reasons. The main concern has been that owners would overcharge tenants for sub-metering services. States that have banned sub-metering include California, New Jersey, Massachusetts, and New York (Allen et al., 2007; NJAA, 2005; Cross, 1996).

5. As we later discuss, we obtain data on contract "switchers" in the post-2013 period, where switchers are defined as firms located in buildings that changed their contract type from owner-paid to tenant-paid utilities, or vice versa. Altogether 65 firms were located in one of these buildings.

6. We use the terms customers, firms, and accounts interchangeably in this study.

7. The 'Industry' category refers to NAICS codes in the Construction, Manufacturing, and Mining sectors.

Table 1: Summary statistics and covariate balance in full sample

	All Firms								
	Tenant-Paid				Owner-Paid				t-Statistic
	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Min	Max	
kW	28.7	43.7	0	582	34.5	62.0	0	576	-0.39
kWh (000s)	8.0	14.0	0	194	9.3	17.3	0	206	-0.31
Bill (\$)	653	1014	12	15,758	739	1231	12	13,128	-0.29
Bill Length	30.3	1.3	27	33	30.4	1.3	27	33	-0.30
Building S.F. (000s)	57.3	59.4	0.8	389	66.7	93.5	2.3	416	-0.42
Year Built	1974	26	1850	2009	1968	33	1855	1992	0.76
Building Stories	2.6	1.6	1	12	3.4	3.1	1	15	-1.09
Industry	0.12	0.33	0	1	0.10	0.30	0	1	0.26
Trade, Accommodation	0.16	0.36	0	1	0.13	0.33	0	1	0.35
Finance, Real Estate, Management	0.47	0.50	0	1	0.55	0.50	0	1	-0.63
Education, Health, Pub. Admin.	0.18	0.38	0	1	0.17	0.37	0	1	0.11
Entertainment, Recreation, Services	0.07	0.25	0	1	0.05	0.22	0	1	0.35
Northeast	0.27	0.44	0	1	0.31	0.46	0	1	-0.35
Southeast	0.20	0.40	0	1	0.23	0.42	0	1	-0.29
Northwest	0.13	0.33	0	1	0.06	0.23	0	1	1.11
Southwest	0.41	0.49	0	1	0.41	0.49	0	1	0.00
Observations	32,764				6,469				
Firms	903				171				

Notes: The table shows the mean, minimum value, maximum value and standard deviation for the observed covariates, for tenant-paid and owner-paid contracts, respectively. The last column shows the value of the t-statistic for the null hypothesis of equal means between the two contract types. The number of observations in each group in the t-statistic calculation is the number of firms in that group. Asterisks indicate a rejection of the null at the 5 percent level of significance.

Hirschberg (1985), and use peak throughput or peak load. Explicitly, for each firm we measure peak throughput or power (in kW) in a billing month, and then take the mean of peak load across all billing months in the sample.⁸ Peak throughput provides one measure of firm size since it reflects the electric load requirements of electricity-using equipment.⁹ We then assign each firm to a firm size decile based on our mean peak load measure.¹⁰ Table 2 presents the sample summary statistics for firms in the top size decile, which includes 19 owner-paid firms and 91 tenant-paid firms. Mean

8. Peak throughput is the maximum power demand in the top hour of each billing month.

9. Commercial equipment such as air conditioners and heat pumps are typically differentiated by size based on kW capacity. For example, the American Society of Heating, Refrigerating and Air Conditioning Engineers categorizes air conditioners and heat pumps with capacity between 20-40 kW as small to medium commercial, and equipment capacities larger than 40 kW as large commercial (ASHRAE, 2004).

10. Our decision to analyze heterogeneity in the data by decile is guided by the simplicity of dividing the sample into ten sub-samples.

Table 2: Summary statistics and covariate balance in top consumption decile

	Top Decile Firms								t-Statistic
	Tenant-Paid				Owner-Paid				
	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Min	Max	
kW	132.4	71.2	17	582	164.2	120.9	5.0	576.0	1.11
kWh (000s)	40.6	24.1	4	194	44.5	34.1	1.3	206.0	0.47
Bill (\$)	3002	1759	414	15,758	3276	2403	169	13128	0.47
Bill Length	30.4	1.3	27	33	30.4	1.3	27.0	33.0	0.03
Building S.F. (000s)	86.8	79.7	2.0	389	144.9	146.4	22.0	416.0	1.68
Year Built	1978	19	1920	2009	1973	24	1900	1990	0.85
Building Stories	3.0	2.4	1	12	6.1	5.1	1	15	2.61*
Industry	0.22	0.41	0	1	0.18	0.39	0	1	0.40
Trade, Accommodation	0.09	0.28	0	1	0.04	0.20	0	1	0.92
Finance, Real Estate, Management	0.46	0.50	0	1	0.77	0.42	0	1	2.83*
Education, Health, Pub. Admin.	0.09	0.29	0	1	0.00	0.00	0	1	2.96*
Entertainment, Recreation, Services	0.15	0.35	0	1	0.00	0.00	0	1	4.09*
North	0.39	0.49	0	1	0.27	0.44	0	1	1.06
South	0.61	0.49	0	1	0.73	0.44	0	1	1.06
City	0.27	0.45	0	1	0.40	0.49	0	1	1.07
Observations	3,202				703				
Firms	91				19				

Notes: The table shows the mean, minimum value, maximum value and standard deviation for the observed covariates, for tenant-paid and owner-paid contracts, respectively. The last column shows the value of the t-statistic for the null hypothesis of equal means between the two contract types. The number of observations in each group in the t-statistic calculation is the number of firms in that group. Asterisks indicate a rejection of the null at the 5 percent level of significance.

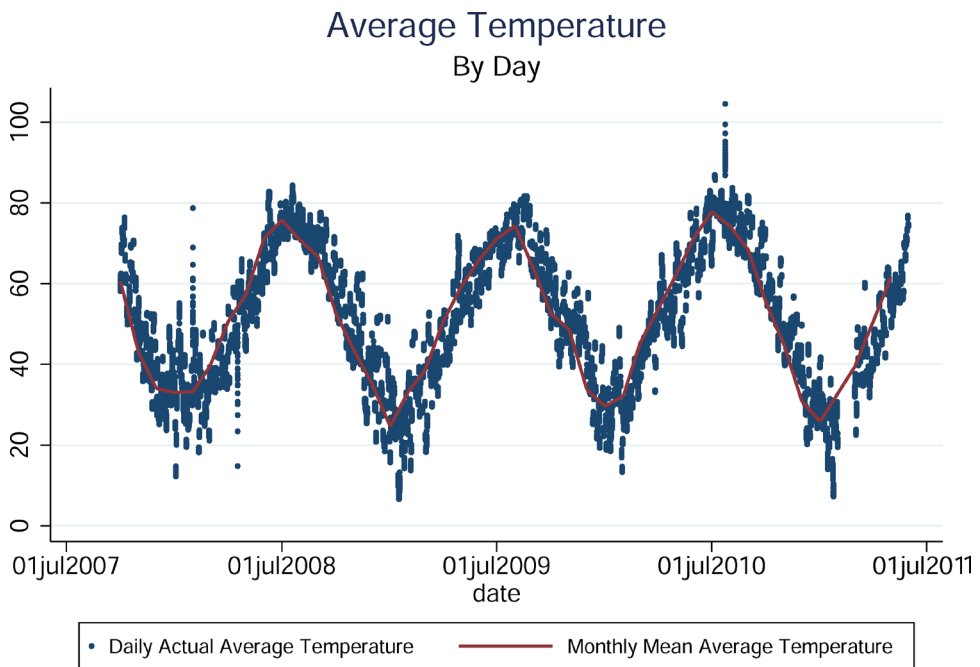
monthly electricity bills in this group are over \$3,000. Buildings in this decile also tend to be located in much larger and somewhat taller buildings compared to the full sample.

Weather is measured as the number of cooling degree days (CDD) and heating degree days (HDD) in a zip code billing-month. To arrive at this observational unit, we begin by using daily temperature data collected from ten local weather stations to construct daily CDD and HDD at each weather station. CDD, which measure demand for space cooling services, are obtained by subtracting 65 from the average Fahrenheit temperature on a given day with temperatures above 65, while HDD, which measure demand for space heating services, are obtained by subtracting the average Fahrenheit temperature on a given day from 65 on days with temperatures below 65. These daily weather station measures are used to compute daily zip code level weather. We use inverse distance weighting relative to zip centroids, and then sum within a billing-month in each zip code to obtain monthly CDD and HDD. Finally, for ease of coefficient interpretation, we divide cumulative CDD and HDD by total days in that billing period to arrive at average daily CDD and HDD by billing month.

This observational unit provides both cross-sectional and temporal variation in weather. One source of cross-sectional variation arises from temperature differences across the 32 zip codes

in UI's service territory. This is made clear in Figure 2, which displays the daily temperature by zip code between October 2007 and May 2011. Despite the relatively small region, there is visible cross-sectional variation in daily temperatures with summer temperatures varying between 5 to 10 degrees across zip codes. Variation in our weather variable also occurs because of differences in billing cycles—which denote the start date and end date of a billing period—across firms. In our sample, there are 16 unique billing cycles, where firm assignment to a billing cycle is based on geography. The staggering of billing cycles throughout a month provides a second source of cross-sectional variation in weather due to the fact that a hot day may be included in different billing “months” for firms on different billing cycles.¹¹

Figure 2: Weather data variation



Notes: Average daily temperature in UI's service territory between October 2007 and May 2011, at the zip code level. Despite the relatively small region, there is visible cross-sectional variation in daily temperatures, with summer temperatures varying between 5 to 10 degrees across zip codes. Temperature variation within a zip code is also possible, due to differences in billing cycles across firms

4. EMPIRICAL FRAMEWORK

In this section, we begin by describing a simple levels comparison of electricity use across firms on owner- and tenant-paid contracts, and show that this approach will likely lead to biased estimates of the principal-agent problem. Next, we detail the empirical approaches that we deploy, the coefficient estimates that these retrieve, the identifying assumptions upon which our empirical approach hinges, and the robustness tests we implement.

11. The assignment of billing cycles based on geography raises the possibility that they may be correlated with weather and contract type. We investigate this by testing if a systematic relationship between bill cycle and weather exists. A regression of weather on bill cycle shows that the sixteen billing cycles are neither jointly nor individually significant in explaining cooling degree days or heating degree days (Appendix section A.1). Nevertheless, our empirical approach explicitly addresses this concern by conditioning on billing cycle.

4.1 Average Treatment Effects: Levels Comparison

To examine the split incentives problem, we begin by comparing overall electricity use across firms on owner- and tenant-paid contracts conditional on a number of rich time controls using OLS,

$$Y_{it} = \alpha + \beta_1 C_{zt} + \beta_2 H_{zt} + \theta T_i + \eta_i t + \gamma_t + \varepsilon_{it} \quad (5)$$

The outcome variable is the natural log of electricity use for firm i in billing month t . The regressor of interest, T_i , is an indicator variable that takes on a value of 1 if firm i is on a utilities excluded or tenant-paid contract, and 0 if it is on a utilities included or owner-paid contract. The variables C_{zt} and H_{zt} are average daily cooling and heating degree days for a firm assigned to billing month t and located in zip code z . We further condition on billing month fixed effects, denoted by γ_t , and firm-specific time trends $\eta_i t$.

Our coefficient of interest, θ , will reflect the average effect of contract type on monthly electricity use if assignment to a tenant-paid or owner-paid contract is independent of potential outcomes. In our setting, this identifying assumption seems untenable, since the mechanism by which firms and buildings are assigned to contract type is likely correlated with fixed firm or building attributes that also determine electricity use. Tenants may sort into contract type based on electricity use, the elasticity of their electricity demand, or firm-specific attributes. Another possibility is that the presence of sub-meters in a building, and hence the ability for owners to implement tenant-paid contracts, may be co-determined with other fixed building attributes. In our setting, the decision to construct a building with or without sub-meters may coincide with other construction decisions such as insulation or window quality that affect electricity use. For these reasons, buildings and firms on tenant-paid contracts likely differ from those on owner-paid contracts in ways that affect electricity use. Failure to account for selection into contract type may result in a biased estimate of θ .

To empirically explore whether selection on fixed firm and building attributes may confound the estimation of equation (5), we compare firms on owner- and tenant-paid contracts across a number of observables that we hypothesize may be related to contract type. Tables 1 and 2 report mean characteristics for firms on tenant- and owner-paid contracts, as well as the t-statistic associated with the difference in means. Motivated by empirical specifications that focus on the principal-agent problem among all firms and only the largest customers, we present these comparisons for all firms in our sample, Table 1, and firms in the top demand decile, Table 2. As shown in Table 1, the covariates are balanced along the rich set of covariates we observe. However, a comparison of means across the top decile of firm size reveals that firms on owner- and tenant-paid contracts differ in building height and industry type. These balance statistics cast doubt on an empirical approach that relies on a levels comparison in electricity use across firms on different contracts, and lead us to forgo the formal estimation of equation (5).

4.2 Average Treatment Effects: Temperature Gradient

We propose an empirical approach that controls for the possibility that firms and buildings on owner- and tenant-paid contracts may be systematically different in fixed attributes that also affect electricity use. We begin with the hypothesis that if a split incentives problem exists, then it should be observed in differences in cooling across owner- and tenant-paid contracts.¹² We test this

12. Most Connecticut commercial customers heat their units with natural gas or fuel oil rather than electricity (EIA, 2012), leading us to hypothesize that electricity use will be most responsive to weather conditions in the summer months, when air-conditioning use is high.

hypothesis by evaluating how electricity use differs in response to a 1 CDD increase across firms on an owner- versus tenant-paid contract, controlling for firm fixed effects and weather.

To evaluate the differential effect of a CDD on electricity use across contract type, we estimate a fixed effects model,

$$Y_{it} = \beta_1 C_{zt} + \beta_2 H_{zt} + \theta_1 T_i \times C_{zt} + \theta_2 T_i \times H_{zt} + L_t + \eta_i t + \gamma_i + \gamma_i + \varepsilon_{it} \quad (6)$$

In this specification, the indicator variable for whether tenant i pays its own electric bill is interacted with each of the weather variables, $T_i \times C_{zt}$ and $T_i \times H_{zt}$. Importantly, this estimating equation conditions on account fixed effects γ_i . This allows us to control for all fixed firm and building characteristics including those that affect electricity use and may systematically differ across contract type. We also condition on bill length, L_t , defined as the number of days in a billing month, to account for differences in weather attributable to variation in bill length across billing months.

The coefficient, θ , reflects the differential effect of temperature increases on electricity use across firms on tenant- and owner-paid contracts. A natural interpretation of θ_1 is the change in demand for air conditioning among tenant-paid contract firms relative to owner-paid firms, in response to warmer temperatures, holding constant the existing building stock. To estimate this treatment effect, we exploit variation in CDD generated from the staggering of billing cycles, and compare how a firm on a tenant- versus owner-paid contract responds to this variation netting out fixed firm characteristics. This approach allows us to account for fixed building and firm attributes systematically correlated with contract type and electricity use.

Nevertheless, identification of the treatment effect still rests on a key assumption: the response of electricity use to CDD differs only by unobservables uncorrelated with contract type. When compared to the levels regression in equation (5), the requirements for identification are less onerous. This is because equation (6) allows for selection into contract type based on fixed unobservables. Our empirical approach only breaks down if fixed building attributes that affect electricity use in a temperature-dependent way are also systematically correlated with contract type. In our setting, this would occur if, for example, building age was systematically correlated with contract type, and the electricity response to temperature differed across building vintage. Under this scenario, our estimated treatment effect would capture the effect of a change in demand for air conditioning among firms in tenant-paid contract buildings relative to owner-paid buildings, where differences across buildings would reflect both contract type and other co-determined building attributes.

To examine the plausibility of our main identifying assumption, we augment equation (6) to account for the possibility that building attributes which differ systematically across contract type may also impact electricity use along a temperature gradient. Our main estimating equation thus conditions on interactions between weather and building and firm attributes that (i) differ across contract type and (ii) are primary determinants of electricity use (EIA, 2015),

$$Y_{it} = \beta_1 C_{zt} + \beta_2 H_{zt} + \theta_1 T_i \times C_{zt} + \theta_2 T_i \times H_{zt} + \psi X_i \times [C_{zt}, H_{zt}] + L_t + \eta_i t + \gamma_i + \gamma_i + \varepsilon_{it} \quad (7)$$

The term $\psi X_i \times [C_{zt}, H_{zt}]$, denotes a vector of building and firm attributes interacted with heating and cooling degree days, where X_i includes indicator variables for building type (retail, office, etc.), firm NAICS code, and quartile of building vintage.¹³

13. The building vintage quartile dummies capture the variation in building energy standard adoptions in Connecticut (OSBI, 2018).

Our testable hypothesis is that if building attributes confound the TRG then our coefficient estimate on contract type, θ_1 , will be sensitive to the inclusion of interactions between temperature and building/firm covariates. If the coefficient estimate remains unchanged after conditioning on these interaction terms, then this provides evidence to support our main identifying assumption and the interpretation of our coefficient of interest as the impact of contract type on electricity use. In terms of the conceptual framework introduced in Section 2, a negative estimate for θ_1 in equation 7 would therefore indicate that the intensive margin (price) dominates, whereas a positive coefficient would indicate that the extensive margin (energy efficiency capital) dominates.

4.3 Conditional Average Treatment Effects: Temperature Gradient

A central focus of this paper is whether the size of the split incentives problem varies substantially across firms. One form of heterogeneity in the response to contract type may arise based on customer size, since relatively larger firms are likely to have higher electricity expenditures, and therefore larger bill savings from conservation. To empirically examine this form of heterogeneity, we estimate conditional average treatment effects for firms in different deciles of average peak power, in (kW).¹⁴ To implement this, we augment equation (7) and estimate,

$$Y_{it} = \beta_{1d}(C_{zt} \times \mathbb{1}_{id}) + \beta_{2d}(H_{zt} \times \mathbb{1}_{id}) + \theta_{1d}(T_i \times C_{zt} \times \mathbb{1}_{id}) + \theta_{2d}(T_i \times H_{zt} \times \mathbb{1}_{id}) + \psi_d X_i \times [C_{zt} \times \mathbb{1}_{id}, H_{zt} \times \mathbb{1}_{id}] + L_t + \eta_i t + \gamma_i + \varepsilon_{it} \quad (8)$$

This estimating equation now includes a vector of indicator variables denoted by $\mathbb{1}_{id}$ that are set equal to 1 if tenant i has mean peak power in decile d (i.e. $d = \{1, \dots, 10\}$), and zero otherwise. These indicator variables are interacted with the weather variables, and the treatment effect of interest. This allows us to separately estimate, for each size decile, the differential effect of a CDD on demand for electricity across contract type.

4.4 Robustness

To examine the plausibility of our main identifying assumption, in addition to the attribute-temperature interaction terms discussed in section 4.2, we implement two novel robustness tests. The first makes use of a regulatory change allowing buildings to switch contract type and tests if selection is an empirical concern. Second, to account for the possibility of remaining selection on unobservables, we apply a new technique proposed by Oster (2017) to bound our estimated treatment effects.

Our first robustness test takes advantage of a policy change to sub-metering regulations. Within our sample period, a ban on sub-metering retrofits in Connecticut made selection by customers and building owners along contract type very costly, if not impossible. For example, customers desiring attributes of a centrally-metered building may have preferred to pay their own electricity, and landlords may have preferred to offer tenant-paid energy utilities. However, retrofitting buildings with unit-level electricity meters—a prerequisite for tenant-paid contracting—was not permitted. In 2013, about two years after our sample period ended, this restriction was lifted and landlords were allowed to retrofit buildings with sub-meters.

14. Details of the size decile construction are discussed in Section 3.1. However, it bears repeating that we define firm size by average peak power (in kW) in the top demand hour within a billing month, and that this variable is distinct from our dependent variable, total monthly energy consumption (in kWh).

We use building-level tenancy contract information collected a year and a half after the Connecticut legislative change to assess whether sorting based on energy consumption preferences might have occurred once sub-metering retrofits were allowed. Since the legislative change allowed a more flexible re-matching of tenants into contract type, this presents an opportunity to observe which buildings switched and to directly examine whether controlling for them changes our baseline results.¹⁵ Under the null hypothesis of “no selection,” our estimated treatment effect should be unchanged after conditioning on the identity of firms switching contract types by interacting indicator variables for these “switchers” with CDD and HDD.

Our second test uses a technique proposed by Oster (2017). This method requires the assumption that the relationship between treatment and unobservables can be recovered from the relationship between treatment and observables. If this is the case, movements in the coefficient of interest and R-squared levels from the inclusion of control variables inform us about selection on unobservables. Building on Altonji et al. (2005), Oster (2017) points out that under the plausible assumption that observable controls share covariance properties with unobservable variables, omitted variable bias is proportional to coefficient movements, but only if these movements are scaled by changes in total R-squared. An ideal scenario in this context is one in which the treatment coefficient of interest changes very little as new covariates are added, and the regression R-squared approaches its maximal theoretically possible value, after accounting for measurement error (Gonzalez and Miguel, 2015). In this case, a large R-squared suggests there is little variation remaining to bias the coefficient. The Oster approach yields a range for the bias-adjusted coefficient of interest, or an identified set formed by the treatment effect in the fully controlled regression, and the bias-adjusted effect. We retrieve the Oster bounded set in a post-estimation procedure and present it in our results.¹⁶

5. RESULTS AND DISCUSSION

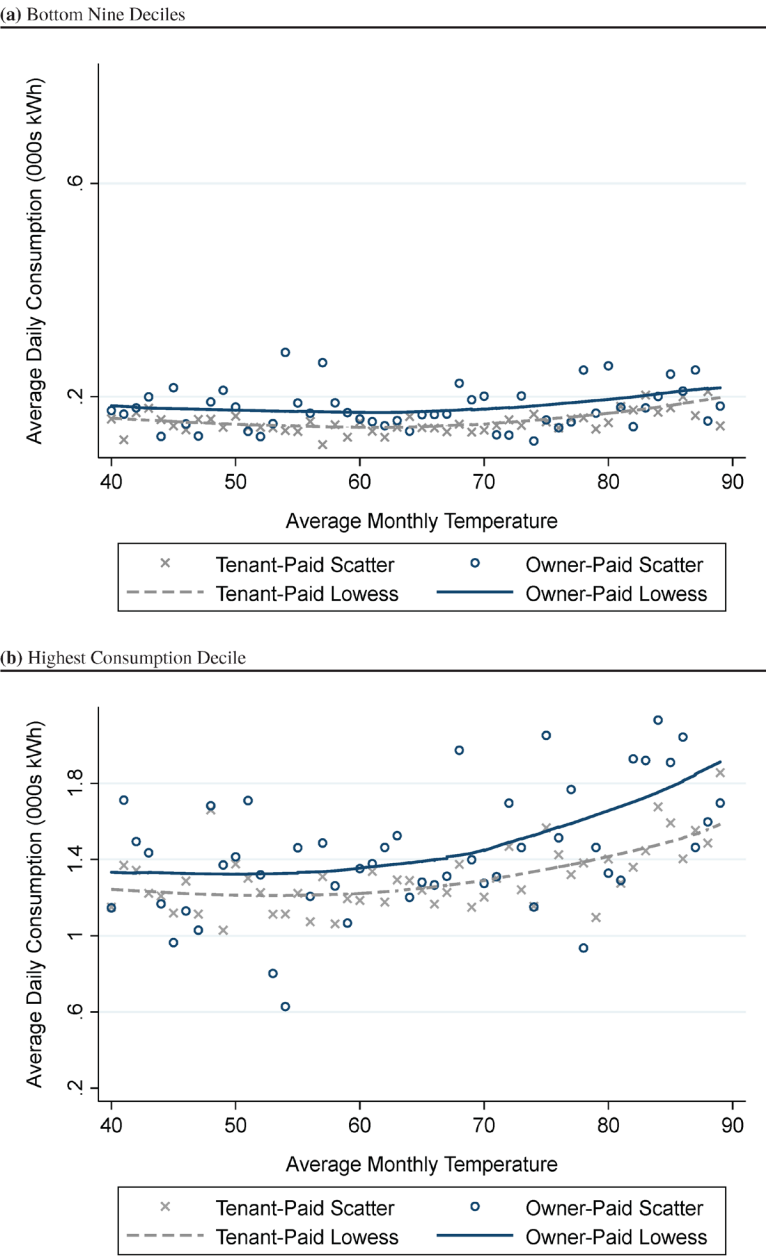
The reduced form relationship between contract type, firm size, temperature and electricity consumption is presented in Figure 3. It plots electricity consumption against average temperature within one-degree bins, across both contract types, for the bottom nine deciles of firm size in panel (a), and the top decile in panel (b). Superimposed on each scatter plot is a lowess fit of consumption on temperature. This figure provides a preview to our formal regression results and points to three interesting patterns of firm behavior. First, as shown in panel (a), on average there is almost no discernible difference in consumption by contract type across the distribution of temperatures in the bottom nine size deciles. Second, in the top size decile, shown in panel (b), we observe a significant divergence in usage across contract types, with firms on owner-paid utility contracts exhibiting higher use. Third, this difference in usage is most pronounced when air-conditioning demand rises. Consumption levels begin to diverge more sharply once temperature increases beyond approximately 65 F, the temperature at which demand for cooling typically begins (EPA, 2014).

Table 3 presents our formal regression results. Column (1) displays estimates from the estimation of equation (6), a regression comparing the differential impact of a weather shock for firms with a tenant-paid contract type relative to an owner-paid contract, controlling for firm and bill-

15. Roughly six percent of customers switched contract types by early 2015, with 34 owners moving to a tenant-paid contract and 31 transitioning to an owner-paid contract. We control for both types of switches in our empirical specifications. Switches to owner-paid contracting were not limited prior to the sub-metering policy change, and there are several reasons why owners may switch to owner-paid contracting, such as metering costs or tenant risk aversion (see Levinson and Niemann (2004) for a more detailed discussion).

16. Appendix Section A.3 provides more detail on the Oster bounds approach.

Figure 3: Consumption by contract type



Notes: Each scatter plot presents monthly electricity consumption against average temperature within 1-degree bins, for the bottom nine decile of firms in panel (a), and the top consumption decile in panel (b). The observations are color-coded by contract type, in both the bottom nine deciles (panel (a)), and the top consumption decile (panel (b)). The solid and dashed lines are a lowess fit of the same data.

ing-month fixed effects and firm-specific time trends. When looking across all firms, we find there is no difference in the effect of weather shocks on consumption across contract type. In the remainder of Table 3, we report results that include tenant-paid contract interactions with CDD and HDD for each size decile. Column (2) reports results from the estimation of the conditional average treatment effects analog of equation (6). Columns (3) to (5) examine the robustness of this result to potential

Table 3: Split incentive effect by consumption decile

Dependent variable:	Log Usage					Log Bill
	(1)	(2)	(3)	(4)	(5)	(6)
Tenant x CDD	-0.00001 (0.00009)					
Tenant x CDD (10th Dec.)		-0.013** (0.006)	-0.015*** (0.006)	-0.014** (0.006)	-0.014** (0.006)	-0.012** (0.005)
Tenant x CDD (9th Dec.)		0.001 (0.009)	0.004 (0.010)	0.005 (0.009)	0.005 (0.009)	0.004 (0.009)
Tenant x CDD (8th Dec.)		-0.000 (0.007)	0.004 (0.007)	0.004 (0.007)	0.004 (0.007)	0.002 (0.005)
Tenant x CDD (7th Dec.)		-0.004 (0.007)	-0.002 (0.008)	0.003 (0.007)	0.003 (0.007)	0.001 (0.005)
Tenant x CDD (6th Dec.)		0.010 (0.008)	0.013* (0.007)	0.010 (0.007)	0.010 (0.007)	0.009* (0.005)
Tenant x CDD (5th Dec.)		0.003 (0.007)	0.004 (0.008)	0.005 (0.007)	0.005 (0.007)	0.004 (0.005)
Tenant x CDD (4th Dec.)		0.009 (0.011)	0.012 (0.011)	0.012 (0.010)	0.012 (0.010)	0.009 (0.006)
Tenant x CDD (3rd Dec.)		-0.017 (0.014)	-0.017 (0.014)	-0.012 (0.013)	-0.012 (0.013)	-0.006 (0.008)
Tenant x CDD (2nd Dec.)		0.005 (0.010)	0.005 (0.010)	0.007 (0.009)	0.007 (0.009)	0.006 (0.005)
Tenant x CDD (1st Dec.)		-0.004 (0.013)	-0.003 (0.014)	-0.003 (0.013)	-0.003 (0.013)	-0.003 (0.007)
Account & Time F.E.s, Acct. Trend	YES	YES	YES	YES	YES	YES
Characteristics Interactions	NO	NO	YES	YES	YES	YES
Characteristics Interactions w/ Year-Built	NO	NO	NO	YES	YES	YES
Switchers Controls	NO	NO	NO	NO	YES	YES
Observations	39,233	39,233	39,233	39,233	39,233	39,233
Accounts	1,074	1,074	1,074	1,074	1,074	1,074
R-squared (within)	0.08	0.09	0.10	0.11	0.11	0.27

Notes: The dependent variable in columns (1)–(5) is the natural log of electricity use in a billing month, and in column (6) it is the natural log of the electricity bill in a billing month. Column (1) presents results without decile interactions, and columns (2)–(6) include results across size deciles. Additional controls included in all regressions are cooling degree days, heating degree days, and heating degree days interacted with contract type. Column (3) further conditions on cooling and heating degree days interacted with building type and NAICS code dummies. Column (4) adds interactions of quartile of year-built with cooling and heating degree days. Column (5) also includes switchers dummies interacted with cooling and heating degree days. Standard errors clustered at the building level are in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

confounding factors by reporting results from the estimation of equation (8). Column (3) conditions on the interaction of CDD and HDD with building and industry type; column (4) adds interactions of CDD and HDD with building vintage quartiles; and column (5) adds controls for the differential effect of temperature shocks among the contract switchers described in Section 4.4.

Our results indicate that a split incentives problem leads to overconsumption of energy among the largest decile of firms. Focusing on our preferred specification in column (5), we find that a tenant-paid contract leads to about a 1.4 percent decrease in kWh per average daily CDD for the top decile of electricity consumers. This translates into about a 3 percent decrease in electricity use among the top size decile of customers. In contrast, contract type does not statistically impact con-

sumption decisions for the other 90 percent of commercial firms. This large divergence in response to contract type based on firm size points to a first source of heterogeneity in response to treatment, and potentially large savings from the targeted deployment of a policy instrument.

A second source of heterogeneity results from seasonal variation in the treatment effect. We find that the split incentive can lead to significant increases in electricity use but only during the hot summer months. In the summer months, switching from an owner to a tenant-paid contract would reduce monthly electricity consumption by up to 14 percent. The summer response is consistent with a framework in which demand for electric air conditioning during these hot months drives the divergence in the TRG across owner- and tenant-paid contracts.¹⁷

Though contract type only influences electricity choices for a narrow set of customers during a concentrated period of time, restructuring contract type has meaningful implications for aggregate electricity usage. This is because the responsive firms are the largest electricity consumers and are quite sensitive to hot temperatures. Our results suggest that a policy which switched the largest decile of electricity consuming firms in our sample from an owner to tenant-paid contract would result in annual electricity savings per firm of roughly 19,000 kWh. Comparing these savings to the total quantity of electricity consumed by all commercial firms in our sample, we find that this policy change would lead to a 1.4 percent reduction in total electricity use.

We also estimate the effect of contract type on electricity expenditure by estimating our preferred conditional average treatment effects specification with log monthly bill as the dependent variable; results are shown in column (6) of Table 3. For the top decile of electricity consumers, the estimated treatment effect is a 1.2 percent decrease in the monthly bill per average daily CDD. The value of total bill savings among these high consumers is approximately \$310 per summer month. On average, this represents a 10 percent reduction in electricity expenditure.

5.1 Robustness Results

One of the robustness tests discussed in Section 4.4 is to use tenancy contract information collected a year and a half after a Connecticut legislative change to assess whether sorting based on energy consumption preferences might have occurred once sub-metering retrofits were allowed. Under the null hypothesis of “no selection,” our estimated treatment effect should be unchanged after conditioning on the identity of firms switching contract types by interacting indicator variables for these “switchers” with CDD and HDD.

This is indeed what we observe. As shown in column (5) of Table 3, including the switchers controls has no significant effect on our estimated coefficient relative to column (4), which does not include the switchers controls. In addition, since columns (3) to (5) of Table 3 include building and firm attributes interacted with heating and cooling degree days, as discussed in Section 4.2, the insensitivity of the results to these provides evidence to support our main identifying assumption and the interpretation of our coefficient of interest as the impact of contract type on electricity use.

To further gauge the robustness of our results to potential selection on unobservables, we apply a second robustness test: the bounds analysis proposed by Oster (2017). We make an equal selection assumption, which implies that any residual omitted variable bias is a function of: (i) the treatment coefficient before and after the inclusion of covariates; (ii) R-squared values before and after the inclusion of covariates; and (iii) the maximum theoretically possible R-squared, namely

17. The coefficients on HDD (not reported) are not statistically significant. Since most firms in Connecticut use natural gas or fuel oil for heating, this is not surprising.

from a regression on consumption and all possible observable and unobservable controls.¹⁸ Given our rich set of controls, the equal selection assumption is likely conservative, as it assumes that any remaining unobservables are as important as the observables in explaining the treatment (Oster, 2017; Altonji et al., 2005).

Table 4 reports the identified set estimates from two different specifications with log usage and log bill as the dependent variables, respectively, corresponding to the fully controlled specifications reported in columns (5) and (6) of Table 3.¹⁹ As shown in this table, we continue to detect a split incentives effect after accounting for any remaining selection on unobservables. A tenant-paid contract induces at minimum monthly electricity and bill savings of 0.7 and 0.6 percent per CDD, respectively. This range implies annual firm-level bill savings of between \$677 to \$1,265.

Table 4: Oster bounds for monthly usage and bill

Log Usage	
Identified Set Estimate	
Lower Bound	-0.014
Upper Bound	-0.007
Log Bill	
Identified Set Estimate	
Lower Bound	-0.012
Upper Bound	-0.006

Notes: The Oster bounds present an identified set of treatment effect coefficients (interpreted as savings per average daily CDD) by accounting for residual omitted variable bias through an equal selection assumption. The omitted variable bias is assumed to be a function of the treatment coefficient and R-squared values before after the inclusion of covariates, as well as the maximum theoretically possible R-squared, namely from a regression on consumption and all possible observable and unobservable controls.

We present further robustness checks in Appendix section A.3. As shown in Table A2, we also include building story interactions with cooling and heating degree days; the results are qualitatively unchanged and the point estimate on our variable of interest increases. Table A2 also shows that our treatment effect is not sensitive to the functional form of the building characteristic controls. Section A.2 also addresses concerns that electricity use may be correlated with mean peak load (the size decile), which could bias our coefficient of interest. To test this concern, we estimate decile-by-decile regressions, thereby eliminating any potential correlation between the decile dummy variable and electricity use. As shown in Table A3, the top decile results are virtually unchanged in the decile-by-decile results.

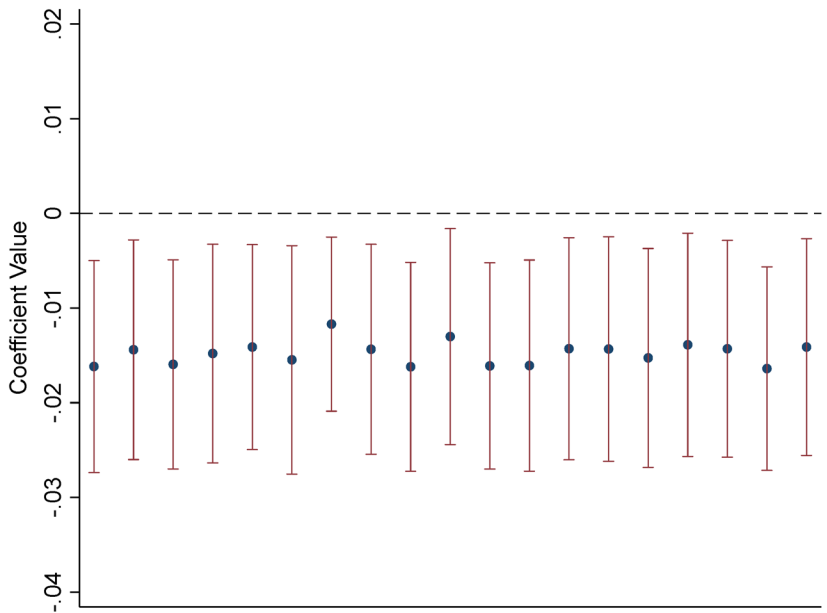
It is possible that there are some residual unobserved attributes of the 19 firms that are correlated with the (true) treatment effect. Such unobservables would have to be time-varying due to the presence of tenant fixed effects, and must also be orthogonal to building type, NAICS code, building vintage, and number of stories in order to introduce bias.

18. Further details on the Oster approach are provided in Appendix Section A.2.

19. The maximum theoretically possible R-squared may be less than 1 if there is measurement error. These set estimates assume that the maximum possible R^2 is 0.98, given the estimated 2 percent measurement error in electricity meter readings (Dong et al., 2005; Reddy et al., 1997).

Finally, since our need to lean on a small sample of top decile firms raises the possibility that our results are driven by outliers, we assess whether the exclusion of any single observation negates our main result. Figure 4 presents coefficient estimates and confidence intervals on the $\text{Tenant} \times \text{CDD}$ interaction term in the 10th decile, obtained by successively dropping one of the 19 owner-paid firms and estimating equation (8). As shown, our results are robust to the exclusion of any of the largest 19 firms on owner-paid contracts.²⁰

Figure 4: Robustness to outliers



Notes: To assess whether outlier observations from one of the 19 owner-paid firms in the top size decile are driving the results, this figure presents coefficient estimates obtained by successively dropping one of the 19 firms. The blue dots represent the point estimates, and the capped lines are the 95% confidence intervals.

5.2 Generalizability

There are roughly 18 million commercial electricity customers in the U.S. and 5.6 million commercial buildings (EIA, 2017; EIA, 2012). In this section, we explore the similarity of the sub-population under study here to the full population of commercial sector tenanted buildings in the U.S. Understanding if our estimates apply to the broader population of large commercial users provides insights into the potential energy savings from restructuring electricity contracts from owner-to tenant-paid. To demonstrate the broader relevance of our results, we proceed in three steps. First, we make use of a representative data set of national commercial building attributes to show that, along important observables, the data source used in our analysis is representative of building attributes throughout the U.S. Second, we focus exclusively on the database used in our analysis, and illustrate that the distribution of attributes for commercial buildings in Connecticut is similar to those in the broader U.S. Third, we compare contract types and energy intensity in commercial

20. The results are also robust to excluding any of the largest 91 tenant-paid firms. These results are not reported here but are available from the authors by request.

buildings in Connecticut to those across the U.S. We use these contract type statistics in Section 5.3 to estimate the energy savings implied by our treatment effect.

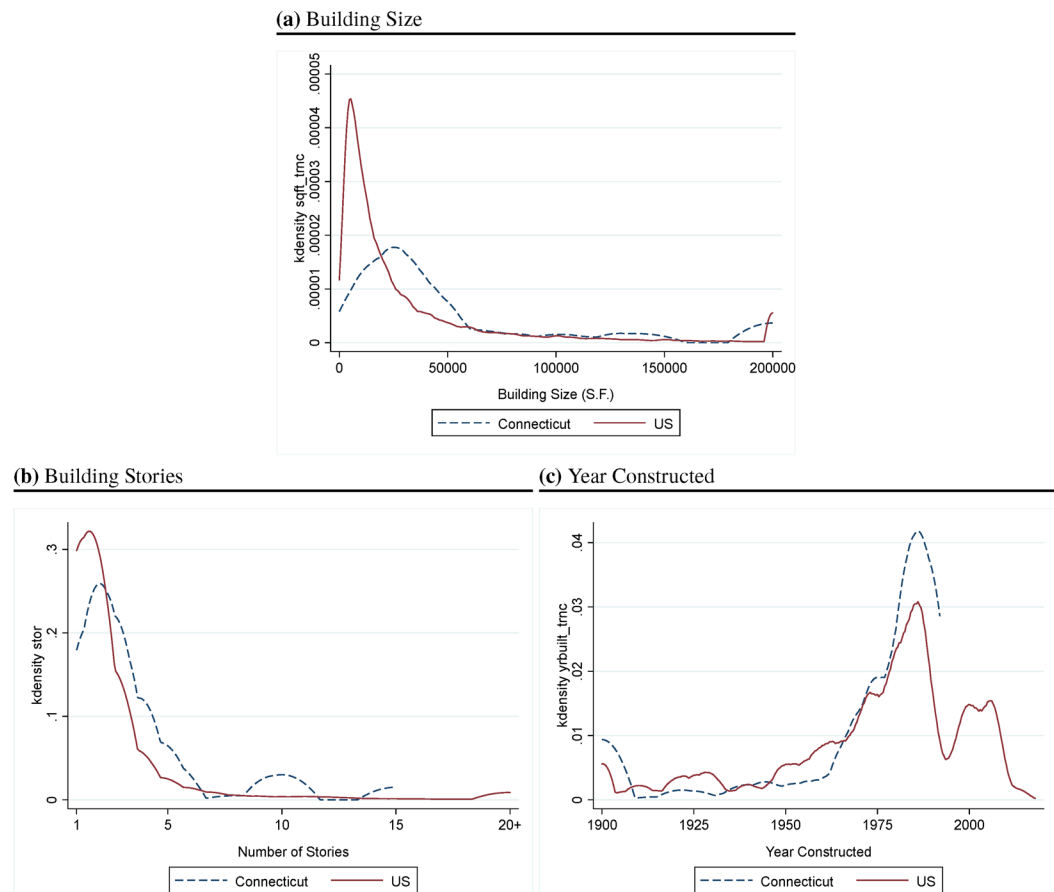
In the first step, we demonstrate that the building database used in our analysis is a representative sample of building attributes in the U.S. Our empirical sample uses data on contract type and building attributes collected from the CoStar group. An advantage of the data collected by the CoStar group is that it includes buildings throughout the U.S., totaling about 97 percent of tenanted buildings. We compare three important building characteristics in the CoStar dataset—building height, age and size—to the Energy Information Administration’s Commercial Building Energy Consumption Survey (CBECS), a nationally representative data set on attributes in both owner and tenant occupied commercial buildings. The CBECS and CoStar datasets are very similar in building height and vintage. While the average CoStar building is larger than the CBECS average, this may be representative of the larger size of leased buildings compared to owner-occupied buildings (EIA, 2012). These similarities in observables, along with the fact that the CoStar database is reflective of leased commercial buildings in the U.S., lend confidence to the national representativeness of the CoStar data.

Second, we show that within the CoStar data there is strong overlapping support in the distributions of measurable building characteristics between Connecticut and the rest of the United States. The overlapping support of building characteristics can be seen in Figure 5. Ideally, we would compare attributes of buildings in the top 10th percentile of electricity usage in Connecticut to those in the U.S. This is not feasible since CoStar does not collect electricity use as a variable. Instead we display the full distribution for both Connecticut and the U.S. of building attributes that we hypothesize are highly correlated with electricity use: square feet, number of stories, and year of construction. For all three variables, significant overlap exists, despite some apparent differences (e.g. Connecticut has a lower proportion of very small buildings). As we discuss below, differences between the Connecticut sample and the broader population imply that the commercial split incentives problem is potentially even larger in the rest of the U.S. than in Connecticut.

Finally, comparing the composition of contract types and energy intensity in Connecticut to the rest of the U.S. leads to the conclusion that the split incentive problem we identify is likely at least as large outside of Connecticut as it is within Connecticut. Approximately 34 percent of commercial, non-government floorspace in New England is leased, as compared to 39 percent nationwide (EIA, 2012). The CoStar database reports contract type for commercial lessees nationwide, differentiating between contracts that transmit price incentives to tenants and those that do not. In our Connecticut sample, about 15 percent of commercial lessees are on owner-paid contracts, as compared to 25 percent nationwide.²¹ With respect to energy intensity, New England is the least energy-intense region in the nation when measured by kWh per square foot of commercial building space (EIA, 2012). When we condition on buildings in which owners pay for electricity, New England is still well below the national average: 11.6 kWh per square foot in New England versus 14.4 nationwide.

Proportionally, less commercial floorspace is rented in New England than nationwide; a higher proportion of commercial renters are on owner-paid contracts in the rest of the U.S.; and the

21. The nationwide figure is even larger if we include contracts with a prorated utility payment for all building occupants, whereby tenants pay a weighted average of the building’s utility bill based on the square feet occupied. In this contractual arrangement, tenants do not pay for the marginal cost of their energy use and large consumers benefit by paying less than their share of utilities. Conservatively, we categorize these as ‘owner-paid’ in our paper, though only about 3 percent of tenants are on a prorated contract in our sample. Nationwide, about 20 percent of tenant contracts include a prorated utility payment. In Section 5.3 we treat these figures under the most conservative assumptions.

Figure 5: Support of building characteristics in Connecticut vs. U.S.

Notes: The Figure shows the overlapping support of building size, stories and year of construction for Connecticut and U.S. buildings.

energy intensity per commercial square foot is higher in regions outside of New England. Thus the magnitude of the potential split incentives problem in the commercial segment is likely to be larger per square foot of building space in the rest of the country than it is in Connecticut.²²

5.3 Quantifying Benefits from Aligning Split Incentives

Under certain plausible assumptions, addressing the commercial split incentives problem for the largest ten percent of commercial firms nationwide has relatively high benefit-to-cost, and would produce energy savings roughly three times larger than those achieved from restructuring rental contracts for all residential users who don't pay their utilities. As detailed in Appendix A.4, using data on the costs of sub-metering, we estimate the payback period and cost effectiveness from sub-metering individual units and shifting to a tenant-paid contract. We find the payback period is less than one year and the cost effectiveness is 3.3 cents per kWh after the first year, 1.6 cents per

22. While we have done our best to assess the representativeness of our sample using measured building characteristics, there remains the possibility that we are missing some remaining time-varying, unmeasured unique aspect of the 19 owner-paid firms, which we use to identify our effect.

kWh after two years, and 1.1 cents after 3 years, assuming the annual electricity savings persist at the same level. As laid out in Appendix A.5, the energy savings from converting the top size decile of firms nationwide from owner- to tenant-paid contracts is 411 gigawatt-hours per year, and amounts to 289 percent of the residential sector analog.²³

We also calculate the reduction in external damages from tenant-paid contracts in our sample. In Appendix section A.6, we convert the estimated energy savings into avoided CO₂ and PM_{2.5} emissions, and then monetize the reduction in emissions. Estimated external benefits and the value of energy savings per firm are presented in Table 5. As shown in columns (1) and (2), the per firm value of avoided damages ranges from \$102 to \$204. In columns (3) and (4) we add to this the estimated bill savings and report an annual firm-level social benefit of switching from an owner- to tenant-paid contract ranging between \$779 and \$1,469. Finally, in columns (5) and (6) we measure the value of the energy savings using the avoided marginal cost of electricity in place of bill savings.²⁴ Total social benefits using avoided marginal costs are between \$676 and \$1,346. Given that the average cost of a sub-meter is \$625, sub-metering retrofits are likely net beneficial from a social perspective.

Table 5: External benefits and the value of energy savings per firm

	<u>External Benefits</u>		<u>External + Value of Savings (Billed)</u>		<u>External + Value of Savings (Marginal Cost)</u>	
	Low \$	High \$	Low \$	High \$	Low \$	High \$
	(1)	(2)	(3)	(4)	(5)	(6)
PM_{2.5}	0.38	0.76	677	1266	574	1143
CO₂	101.95	202.91	779	1468	676	1345
Total	102.33	203.67	779	1469	676	1346

Notes: External Benefits measure the annual per-firm reduction in pollution damages from lower electricity consumption. External + Value of Savings (Billed) measures the sum of the external benefits and the value of the bill savings from contract type, which are the annual bill savings noted in the text (\$677-\$1487). External + Value of Savings (Marginal Cost) uses the average hourly locational marginal price in Connecticut over the sample period, of \$59.42, to value the energy savings. The low and high values are derived from the Oster identified set estimates for electricity savings, discussed in the text.

One caveat in the interpretation of these benefit calculations is that we focus exclusively on the net effect of a switch from owner to tenant-paid contracts on the TRG. The possibility remains that such a switch can lead to a countervailing underinvestment in energy efficiency capital that may impact the level of electricity use. While many building-level energy efficiency investment choices available to landlords, such as insulation or air-conditioner models, affect the TRG, we cannot rule out that the level effects may mitigate the magnitude of the energy savings implied by our estimates, implying a potential role for complementary building energy standards.

5.4 Bill Savings and the Non-Response of Most Commercial Firms

While we estimate that contract type has a sizable effect on electricity use for the largest firms, one unanswered question is why the remaining 90 percent of commercial firms do not re-

23. Under extremely conservative assumptions detailed in Appendix [residential]A.5, the savings are still 125 percent of the residential sector analog.

24. We use this approach to net out fixed costs. Fixed costs are not avoided costs in this setting, since they will be recovered by the utility from other customers under the cost-plus regulatory structure in Connecticut. Our measure of avoided marginal cost is the average hourly locational marginal price for Connecticut over the sample period, \$59.42. Our data source is the New England Independent System Operator (NE-ISO), www.iso-ne.com.

spond to contract type. In our view, the most likely explanation is that the net benefit of decreasing electricity consumption for these customers is negative. This is consistent with recent research that documents negative realized net benefits from energy efficiency investments (Hassett and Metcalf, 1999; Fowle et al., 2015). Consider the electricity choices of an office building, the sector that makes up the largest share of buildings in our sample. Overcooling and overheating are common in office buildings. Derrible and Reeder (2015) suggest that overcooling increases electricity consumption by 8 percent per year. For the average top decile firm in our sample, this would translate into roughly 40,000 of wasted kWh that could be eliminated by managerial oversight or behavioral change. Such actions would translate into \$3,200 in bill savings. By comparison, 8 percent of usage for the average firm in the next-largest (ninth) decile translates into 15,000 kWh, or \$1,200 in potential savings. Accessing these savings would likely require incurring a fixed cost, such as allocating attention of a manager or engineer to monitor and adjust air conditioner and chiller operation. For all but the largest firms, the cost of avoiding overcooling may well exceed the reduction in expenditure from wasted energy.

Other explanations could also account for the lack of a treatment effect across most firms. One possibility is that a tenant-paid contract does induce some conservation behavior among the bottom 9 deciles, but that landlord-side underinvestment in building-level energy efficiency cancels any of the consumption savings from tenants paying their own utility bills. Another possibility is (potentially rational) inattention leading to unresponsiveness among commercial firms (Jesoe and Rapson, 2015). Comparing the \$677 to \$1,265 annual bill saving from a tenant-paid contract to the average commercial unit size in Connecticut, 14,000 square feet, implies an average annual bill saving of 4.8 to 9 cents per square foot. This represents about 0.2 percent of the average annual revenues per square foot in office and retail industries and highlights that the savings smaller firms forgo likely represent a small share of their annual sales. After accounting for the time and effort required to calculate the energy savings from different energy efficiency investments, smaller firms may be rationally inattentive to potential energy savings (Sallee, 2014).

6. CONCLUSION

We measure the “split incentive” effects of tenancy contract type using a unique empirical setting and dataset of tenancy contracts and electricity use among commercial sector clients. Our empirical framework compares how temperature shocks impact electricity consumption across firms on owner- and tenant-paid contracts. Importantly, it helps us to overcome the well-known empirical challenge of separately identifying the split incentives problem from selection on fixed attributes.

Our approach consists of three steps to probe and address the main identification challenge: selection on unobservables that affect electricity use along a temperature gradient. We allow for a heterogeneous TRG along several dimensions by including interactions between temperature and building attributes that may be correlated with energy consumption, testing for selection by taking advantage of a state-level change in metering regulations, and accounting for any potential remaining correlations between unobservable characteristics and the treatment using the Oster (2017) identified set approach.

Our results indicate heterogeneous returns to a tenant-paid contract, with a negative and significant effect of the tenant-paid contract type on consumption observable only in the top decile of firm size. The intensive margin effect dominates among the largest firms. This heterogeneous response is consistent with a setting in which the bill savings from changing consumption do

not cover the adjustment costs for small firms, and is in line with recent evidence from the residential sector (Elinder et al., 2017).

This implies a potential policy case for encouraging tenant-paid energy contracting among large commercial and industrial customers. For the largest decile of firms, we find that firms who pay their own utility bills consume about 3 percent less electricity annually than tenants whose utility bills are bundled into rents, and save between \$677 and \$1,265 on their annual electricity bills. These reductions lead to a 1.4 percent saving in total electricity consumed by all firms in our sample. The payback period from sub-metering and switching to a tenant-paid contract is less than one year when considering only the temperature response gradient.

However, a policy mandating such a switch may also impact the level of electricity use via owner investment decisions. Mandating tenant-paid contracts could lead owners to invest less in energy efficient capital, which would affect energy consumption levels. Consider a world in which landlords invest more in energy efficiency when they bear the costs of their tenant's electricity use. Since owner-paid buildings under this line of reasoning already enjoy a high level of energy efficient capital, if they were to become tenant-paid, not much would change since energy efficiency is both durable and in most cases not transportable from one building to another. There's little reason to expect landlords to spend money to *un-install* past energy-efficiency investments. As a result, only going-forward investments would be affected by this contract switch. There is little evidence to draw upon that would allow us to compare the magnitude of these investment effects to the gradient effects measured in this paper. However, while these may well generate sizeable effects over the long term, they could also be small and may be outweighed by the gradient effects.

Despite the robustness of our results, there is still a need for complementary evidence via additional research. Statistical power is a concern throughout empirical science (Ioannidis, 2005), and studies with samples of this size may be at risk of being underpowered. Nonetheless, the difficulty of developing and deploying a research design to test the split incentives problem among commercial electricity customers makes ours one of the very first contributions to quantify the commercial split incentives effect on energy consumption.

Acknowledgements

We thank Hassan Benchechrone, Laura Grant, Elisabeth Isaksen, Justin Leroux, Leslie Martin, Pierre Merel, Nic Rivers, Michael Roberts, and four anonymous referees for helpful comments and suggestions. The paper also benefited from participant comments from presentations at McGill University, the Canadian Resource and Environmental Economists' Study Group, the Association of Environmental and Resource Economists annual conference in Breckenridge, Colorado, and the USC Workshop on Microeconomic Impacts of Climate Change. We gratefully acknowledge financial support from the Social Sciences and Humanities Research Council of Canada. Any errors are our own.

REFERENCES

- AHRI (2017). "Zone Control System Technology." Air-Conditioning, Heating and Refrigeration Institute.
- Aigner, Dennis J. and J. Hirschberg (1985). "Commercial/Industrial Customer Response to Time-of-Use Electricity Prices: Some Experimental Results." *RAND Journal of Economics* 15: 341–355. <https://doi.org/10.2307/2555562>.
- Allcott, Hunt (2011). "Social Norms and Energy Conservation." *Journal of Public Economics* 95: 1082–1095. <https://doi.org/10.1016/j.jpubeco.2011.03.003>.
- Allen, P.V., P.C. Lacourciere, and R.M. Shapiro (2007). "Submetering of Electricity for Commercial Buildings." November, Thellen LLP, published by Mondaq Ltd.
- Altonji, Joseph G., Todd Elder, and Christopher R. Taber (2005). "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113: 151–184. <https://doi.org/10.1086/426036>.
- Anderson, Soren T. and Richard G. Newell (2004). "Information programs for technology adoption: the case of energy-efficiency audits." *Resource and Energy Economics* 26: 27–50. <https://doi.org/10.1016/j.reseneeco.2003.07.001>.
- ASHRAE (2004). "HVAC Systems and Equipment: Chapter 45-Unitary Air Conditioners and Heat Pumps." American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

- Blumstein, Carl, Betsy Krieg, Lee Schipper, and Carl York (1980). "Overcoming Social and Institutional Barriers to Energy Conservation." *Energy* 5: 355–371. [https://doi.org/10.1016/0360-5442\(80\)90036-5](https://doi.org/10.1016/0360-5442(80)90036-5).
- Connors, S., K. Martin, M. Adams, E. Kern, and B. Asiamah-Adjei (2005). "Emissions Reductions from Solar Photovoltaic (PV) Systems." MIT Laboratory for Energy and the Environment report 2004-003.
- Cross, Phillip S. (1996). "N.Y. Allows Gas Submetering for C&I Customers." *Fortnightly Magazine*, November 15, 1996.
- CRS (2017). "The Regional Greenhouse Gas Initiative: Lessons Learned and Issues for Congress." Congressional Research Service. Davis, Lucas W. (2012). "Evaluating the Slow Adoption of Energy Efficiency Investments: Are Renters Less Likely to Have Energy Efficient Appliances?" in Don Fullerton and Catherine Wolfram eds. *The Design and Implementation of U.S. Climate Policy*.
- Derrible, Sybil and Matthew Reeder (2015). "The Cost of Over-Cooling Commercial Buildings in the United States." *Energy and Buildings* 108: 304–306. <https://doi.org/10.1016/j.enbuild.2015.09.022>.
- DOE (2017). "The Weatherization Assistance Program: An American Industry." Office of Energy Efficiency and Renewable Energy, Department of Energy.
- Dong, Bing, Siew Eang Lee, and Majid Haji Sapar (2005). "A Holistic Utility Bill Analysis Method for Baseline Whole Commercial Building Energy Consumption in Singapore." *Energy and Buildings* 37: 167–174. <https://doi.org/10.1016/j.enbuild.2004.06.011>.
- EIA (2009). "Residential Energy Consumption Survey." Energy Information Administration.
- EIA (2010). "Weatherization Assistance Program: Know the Facts." February 2010 Short Term Energy Outlook, Energy Information Administration.
- EIA (2012). "Commercial Building Energy Consumption Survey." Energy Information Administration.
- (2015). "A Look at the U.S. Commercial Building Stock: Results from EIA's 2012 Commercial Buildings Energy Consumption Survey (CBECS)." Energy Information Administration.
- EIA (2017). "Electric Power Monthly." Energy Information Administration.
- Elinder, Mikael, Sebastian Escobar, and Ingel Petre (2017). "Consequences of a price incentive on free riding and electric energy consumption." *Proceedings of the National Academy of Sciences*, Vol. 114. <https://doi.org/10.1073/pnas.1615290114>.
- EPA (2014). "Heating and Cooling Degree Days." Environmental Protection Agency, <http://bit.ly/1RywQQL>.
- Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram (2018). "Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program." *Quarterly Journal of Economics* 133: 1597–1644. <https://doi.org/10.1093/qje/qjy005>.
- Gillingham, Kenneth, Matthew Harding, and David Rapson (2012). "Split Incentives in Household Energy Consumption." *The Energy Journal* 33(2): 37–62. <https://doi.org/10.5547/01956574.33.2.3>.
- Gillingham, Kenneth and Karen Palmer (2014). "Bridging the Energy Efficiency Gap: Insights for Policy from Economic Theory and Empirical Analysis." *Review of Environmental Economics and Policy*, Vol. 8. <https://doi.org/10.1093/reep/ret021>.
- Giraudet, Louis-Gaëtan and Dominique Finon (2015). "European experiences with white certificate obligations: A critical review of existing evaluations." *Economics of Energy & Environmental Policy* 4: 113–130. <https://doi.org/10.5547/2160-5890.4.1.lgir>.
- Gonzalez, Felipe and Edward Miguel (2015). "War and Local Collective Action in Sierra Leone: A Comment on the Use of Co-efficient Stability Approaches." *Journal of Public Economics* 128: 30–33. <https://doi.org/10.1016/j.jpubeco.2015.05.004>.
- Hartford Business Journal (2013). "CT approves submetering, fines N.H. landlord. June 10, 2013." retrieved December 2017.
- Hassett, Kevin A. and Gilbert E. Metcalf (1999). "Measuring the Energy Savings from Home Improvement Investments: Evidence from Monthly Billing Data." *The Review of Economics and Statistics* 81: 516–528. <https://doi.org/10.1162/003465399558274>.
- Hoger, Ryan R. (2014). "Proper Applications and Sequences for Constant Volume Pressure Dependent Zoning Systems." March, HVAC for Roof-Top Units, Massachusetts Energy Efficiency Partnership.
- Ioannidis, John (2005). "Why Most Published Research Findings Are False." *PLoS Medicine* 2: 696–701. <https://doi.org/10.1371/journal.pmed.0020124>.
- IWGSCC (2015). "Technical Support Document: Technical Update on the Social Cost of Carbon for Regulatory Analysis Under Executive Order 12866." Interagency Working Group on Social Cost of Carbon, United States Government.
- Jaffe, Adam and Robert Stavins (1994). "Energy Efficiency Investments and Public Policy." *The Energy Journal* 15: 43–65. <https://doi.org/10.5547/ISSN0195-6574-EJ-Vol15-No2-3>.
- Jessoe, Katrina and David Rapson (2015). "Commercial and Industrial Demand Response Under Mandatory Time-of-Use Electricity Pricing." *Journal of Industrial Economics* 63: 397–421. <https://doi.org/10.1111/joie.12082>.
- 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* 114: 1–12. <https://doi.org/10.1016/j.jpubeco.2014.03.003>.

- Krishnamurthy, C.K.B. and Bengt Kriström (2015). “How Large is the Owner-Renter Divide in Energy Efficient Technology? Evidence from an OECD Cross-section.” *The Energy Journal* 36: 85–103. <https://doi.org/10.5547/01956574.36.4.ckri>.
- Levinson, A. and S. Niemann (2004). “Energy use by apartment tenants when landlords pay for utilities.” *Resource and Energy Economics* 26: 51–75. [https://doi.org/10.1016/S0928-7655\(03\)00047-2](https://doi.org/10.1016/S0928-7655(03)00047-2).
- Meyer, Leo A. (2006). “Variable Air Volume Systems.” May, The Air Conditioning, Heating and Refrigeration News.
- Millstein, Don (2008). “Submetering can measure energy use more accurately in shopping centers.” *Retail Construction Magazine*, January/February 2008.
- Muller, Nicholas Z. and Robert Mendelsohn (2007). “Measuring the damages of air pollution in the United States.” *Journal of Environmental Economics and Management* 54: 1–14. <https://doi.org/10.1016/j.jeem.2006.12.002>.
- Murtishaw, S and J Sathaye (2006). “Quantifying the Effect of the Principal-Agent Problem on U.S. Residential Energy Use.” Lawrence Berkeley National Laboratory report LBNL-59773. <https://doi.org/10.2172/913156>.
- Myers, Erica (2015). “Asymmetric Information in Residential Rental Markets: Implications for the Energy Efficiency Gap.” Energy Institute at Haas Working Paper 246.
- NJAA (2005). “Water Conservation and Economy in New Jersey through Sub-Metering of Water in Multi-Family Rental Housing.” New Jersey Apartment Association.
- Novan, Kevin, Aaron Smith, and Tianxia Zhou (2017). “Residential Building Codes Do Save Energy: Evidence From Hourly Smart-Meter Data.” June, Working Paper.
- OSBI (2018). “Office of the State Building Inspector.” <https://bit.ly/2siATNW>. Last accessed May 30, 2018.
- Oster, Emily (2017). “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business Economics and Statistics*. <https://doi.org/10.1080/07350015.2016.1227711>.
- Papineau, Maya (2017). “Energy Efficiency Premiums in Unlabeled Office Buildings.” *The Energy Journal* 38: 195–212. <https://doi.org/10.5547/01956574.38.4.mpap>.
- Pike Research (2012). “Electricity Submeters.” Research Report.
- Reddy, T.A., N.F. Saman, D.E. Claridge, J.S. Haberl, W.D. Turner, and A. Chalifoux (1997). “Baselining Methodology for Facility-Level Monthly Energy Use—Part 1: Theoretical Aspects.” *ASHRAE Trans.*, Vol. 103.
- Rothschild, Susy and Art Diem (2009). “Total, Non-baseload, eGRID Subregion, State? Guidance on the Use of eGRID Output Emission Rates.” E.H. Pechan & Associates and U.S. EPA.
- Sallee, James (2014). “Rational Inattention and Energy Efficiency.” *Journal of Law and Economics* 57: 781–820. <https://doi.org/10.1086/676964>.
- Sanchez, Marla, Carrie Webber, Richard Brown, John Busch, Margaret Pinckard, and Judy Robertson (2007). “Space Heaters, Computers, Cell Phone Chargers: How Plugged In Are Commercial Buildings?.” Lawrence Berkeley National Laboratory Report No. 62397.
- Stavins, Robert N. (2011). “The Problem of the Commons: Still Unsettled after 100 Years.” *American Economic Review* 101: 81–108. <https://doi.org/10.1257/aer.101.1.81>.
- White, Evan (2012). “Utilities in Federally Subsidized Housing.” Goldman School of Public Policy and UC Berkeley Law at Boalt Hall.

APPENDIX

A.1 Bill Cycles and Weather

We assess whether bill cycle is correlated with the TRG across contract type by testing for a systematic relationship between bill cycle and weather. In Table A1, we report the results of a regression of weather on bill cycle. As shown, we find that the sixteen billing cycles are neither jointly nor individually significant in explaining cooling degree days or heating degree days.

A.2 Oster Bounds Details

Our study assesses the effect of contract type T on electricity consumption y , as specified in equation (8). One of our robustness tests implements the Oster (2017) approach. A more detailed explanation of the approach follows here.

Table A1: Bill cycle conditional independence assumption

	(1) CDD	(2) CDD 10th Dec.	(3) HDD	(4) HDD 10th Dec.
Bill Cycle 2	-0.037 (0.466)		0.306 (1.657)	
Bill Cycle 3	-0.131 (0.468)	0.114 (0.814)	0.298 (1.674)	-0.605 (2.429)
Bill Cycle 4	0.189 (0.434)	0.433 (0.557)	-0.394 (1.553)	-2.071 (1.663)
Bill Cycle 5	-0.131 (0.475)	-0.044 (0.713)	0.197 (1.703)	-1.172 (2.139)
Bill Cycle 6	-0.156 (0.527)	0.154 (0.814)	0.111 (1.886)	-1.000 (2.429)
Bill Cycle 7	0.050 (0.454)	0.326 (0.596)	0.416 (1.620)	-1.205 (1.806)
Bill Cycle 8	-0.182 (0.499)	0.184 (0.620)	0.490 (1.775)	-0.594 (1.865)
Bill Cycle 9	0.470 (0.501)	-0.052 (0.632)	-0.509 (1.796)	-0.911 (1.936)
Bill Cycle 10	0.621 (0.509)	-0.196 (0.686)	-1.142 (1.833)	-0.966 (2.160)
Bill Cycle 11	-0.028 (0.487)	0.137 (0.733)	0.646 (1.754)	-1.850 (2.275)
Bill Cycle 12	-0.061 (0.438)	-0.168 (0.579)	0.170 (1.572)	-0.906 (1.783)
Bill Cycle 13	0.130 (0.484)	0.006 (0.686)	-0.785 (1.744)	-2.248 (2.183)
Bill Cycle 14	0.024 (0.448)	-0.151 (0.586)	0.781 (1.600)	-0.751 (1.783)
Bill Cycle 15	0.157 (0.569)	0.190 (0.709)	0.354 (2.003)	0.208 (2.128)
Bill Cycle 16	-0.003 (0.516)	0.162 (0.679)	0.466 (1.852)	-1.849 (2.118)
Constant	2.193*** (0.377)	2.543*** (0.451)	15.408*** (1.358)	16.748*** (1.335)
Observations	2,479	1,051	2,594	1,270
F test for joint significance	0.471	0.218	0.248	0.245

If observables and unobservables have the same explanatory power in y (after taking into consideration any measurement error in y), then the following is a consistent estimator of the effect of T on y :

$$\hat{\theta} = \hat{\theta}^* - (\hat{\theta} - \hat{\theta}^*) \times \frac{R_{max} - R^*}{R^* - R}, \quad (A1)$$

where $\hat{\theta}^*$ and R^* are the coefficient estimate and R^2 from the regression with a full set of controls, in column (5) of Table 3, and $\hat{\theta}$ and R are the coefficient and R^2 from an uncontrolled regression that includes only the coefficient of interest. The R_{max} term represents the R^2 in a regression of y on all possible observable and unobservable controls, which can be less than 1 in there is measurement error in y . The $\hat{\theta}$ is the basis for identifying the upper bound for the Oster bounds presented in Table 4 of the main paper.²⁵

A.3 Additional Robustness Tests

Our estimated treatment effect is not sensitive to alternative specifications, as shown in Table A2. Column (1) is the fully controlled specification from column (5) of Table 3, augmented with stories quartile dummies interacted with cooling and heating degree days. The point estimate increases and remains statistically significant. In columns (2)–(7) we show that the results are not sensitive to the functional form of the building characteristic controls. The point estimate changes very little when the characteristics are included as is or in the form of tertile, quintile or sextile dummies.

Another robustness check we conduct is to estimate decile-by-decile regressions instead of interacting decile dummy variables with the contract type dummy in the full sample, as in equation (8). This is to assess whether consumption in (kWh) may be correlated with average power in the peak hour within a billing month, which we use to construct our size deciles. If that is the case, one might be concerned this correlation could bias the coefficients.²⁶ Estimating ten separate regressions for each decile avoids any potential correlation of the decile dummy variables with consumption.

The results of these regressions are presented in Table A3. They confirm what we find in the full sample results in Table 3 of the paper. In the fully controlled specification, shown in column (4), the treatment effect indicates a 1.3 percent reduction in monthly consumption per average monthly cooling degree day induced by a tenant-paid contract. This is lower by 0.1 percentage points compared to Table 3 in the paper. These decile-by-decile regressions also show that in the sixth size decile a tenant-paid contract is associated with an increase in consumption and monthly bill relative to an owner-paid contract. Based on our conceptual model, this suggests that capital inefficiencies may dominate any benefit from a price signal in this decile. However, the share of the within variation in consumption explained by the different specifications in the sixth decile is among the lowest, which suggests that there is considerable remaining unobserved heterogeneity.

Finally, to allay any concerns regarding a potential relationship between firm deciles and the kWh-temperature gradient, we have created Figure A1 reporting the consumption response to temperature within each decile. The figure reports the coefficients of a decile-by-decile regression of normalized kWh on average daily temperature. The coefficients report values for deciles 1 through 10. As shown, there is no evidence of a monotonic relationship between consumption and tempera-

25. As noted in the main paper, given our rich set of controls the equal selection assumption we make is likely conservative, as it assumes that any remaining unobservables are as important as the observables in explaining the treatment (Oster, 2017; Altonji et al., 2005).

26. The correlation between the size deciles and monthly kWh consumed is 0.58, so while it is positive, it is by no means perfect. There are many firms with high average kW and low kWh and (to a lesser extent) vice versa.

Table A2: Robustness to alternative specifications

Dependent variable:	Log Usage						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tenant x CDD (10th Dec.)	-0.015** (0.007)	-0.014*** (0.004)	-0.016*** (0.004)	-0.017*** (0.005)	-0.017*** (0.005)	-0.017*** (0.005)	-0.016*** (0.005)
Tenant x CDD (9th Dec.)	0.006 (0.008)	0.004 (0.011)	0.001 (0.010)	0.000 (0.010)	0.001 (0.010)	0.001 (0.010)	0.001 (0.010)
Tenant x CDD (8th Dec.)	0.005 (0.007)	-0.001 (0.008)	-0.002 (0.007)	-0.003 (0.007)	-0.001 (0.007)	-0.001 (0.007)	-0.001 (0.007)
Tenant x CDD (7th Dec.)	0.002 (0.007)	-0.007 (0.008)	-0.008 (0.008)	-0.007 (0.008)	-0.007 (0.008)	-0.008 (0.008)	-0.007 (0.008)
Tenant x CDD (6th Dec.)	0.009 (0.007)	0.010 (0.008)	0.009 (0.008)	0.010 (0.008)	0.010 (0.008)	0.009 (0.008)	0.010 (0.008)
Tenant x CDD (5th Dec.)	0.006 (0.007)	0.000 (0.006)	-0.001 (0.006)	-0.001 (0.007)	-0.000 (0.007)	-0.000 (0.007)	-0.000 (0.007)
Tenant x CDD (4th Dec.)	0.012 (0.010)	0.009 (0.010)	0.009 (0.010)	0.009 (0.010)	0.009 (0.010)	0.009 (0.010)	0.009 (0.010)
Tenant x CDD (3rd Dec.)	-0.013 (0.012)	-0.018 (0.014)	-0.019 (0.014)	-0.019 (0.014)	-0.018 (0.015)	-0.019 (0.014)	-0.019 (0.014)
Tenant x CDD (2nd Dec.)	0.006 (0.010)	0.006 (0.009)	0.005 (0.009)	0.006 (0.009)	0.007 (0.009)	0.006 (0.009)	0.006 (0.009)
Tenant x CDD (1st Dec.)	-0.004 (0.013)	-0.007 (0.013)	-0.008 (0.012)	-0.007 (0.013)	-0.007 (0.013)	-0.007 (0.013)	-0.008 (0.013)
Account Fixed Effects	YES	NO	NO	NO	NO	NO	NO
Time F.E.s, Acct. Trend	YES	YES	YES	YES	YES	YES	YES
Switchers Controls	YES	YES	YES	YES	YES	YES	YES
Characteristics Controls	YES	YES	YES	YES	YES	YES	YES
Observations	39,233	39,233	39,233	39,233	39,233	39,233	39,233
Accounts	1,074	1,074	1,074	1,074	1,074	1,074	1,074
R-squared (within)	0.110	0.085	0.085	0.085	0.085	0.085	0.085

Notes: The dependent variable in columns (1)-(7) is the natural log of electricity use in a billing month. Column (1) augments the specification estimated in column (5) of Table 2 to include building stories quartile dummies interacted with cooling and heating degree days. Columns (2)-(7) present specifications without firm fixed effects. Column (2) includes building type and NAICS code dummy variables, year of construction, number of stories and building size in square feet. Column (3) replaces the number of stories with dummy variables for each story. Column (4) includes quartile dummies for year of construction, number of stories, and building size. Columns (5)-(7) includes the same variables in the form of tercile, quintile and sextile dummies, respectively. Additional controls included in all regressions are cooling degree days, heating degree days, and heating degree days interacted with contract type. Standard errors clustered at the building level are in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

ture moving from the lower to upper deciles. For example, the third, fourth and ninth deciles have a higher consumption response to temperature coefficient compared to the top decile.

A.4 Benefit-to-Cost Calculation

Sub-meter costs range from \$250-\$1000 per unit (Pike Research, 2012; White, 2012; Millstein, 2008). Given the average estimated annual bill savings of \$970 and assuming a unit-level sub-meter cost of \$625, the payback period is less than one year, even after allowing for installation

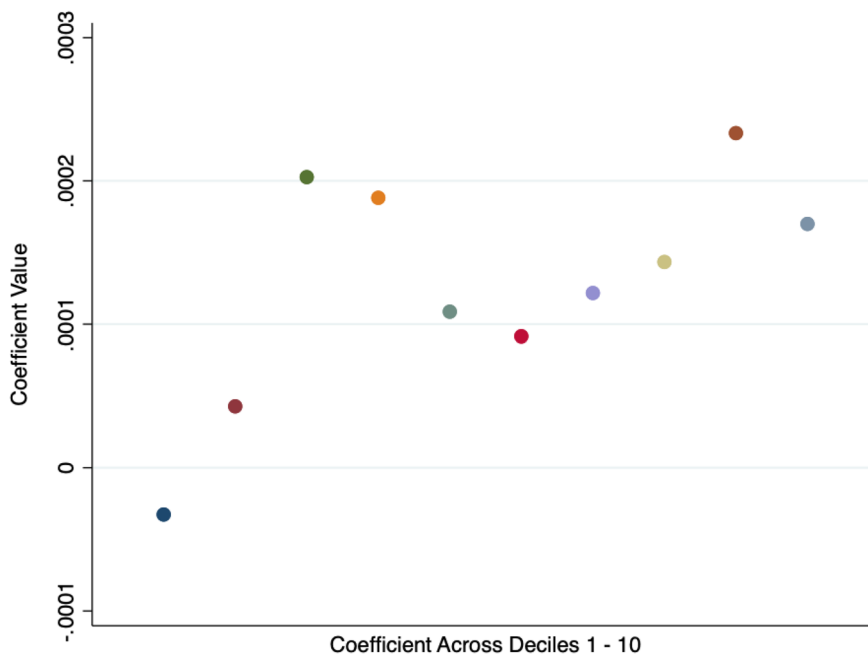
Table A3: Decile-by-decile regressions

Dependent variable:	Usage				Log Bill	
	(1)	(2)	(3)	(4)	(5)	
Tenant x CDD (10th Dec.)	-0.012** (0.006)	-0.013** (0.006)	-0.013** (0.006)	-0.013** (0.006)	-0.012** (0.005)	Observations 3,905
R-squared (within)	0.197	0.273	0.282	0.288	0.433	
Tenant x CDD (9th Dec.)	0.001 (0.010)	-0.000 (0.010)	0.001 (0.008)	0.003 (0.009)	0.003 (0.008)	Observations 3,952
R-squared (within)	0.078	0.157	0.194	0.205	0.324	
Tenant x CDD (8th Dec.)	0.001 (0.007)	0.007 (0.009)	0.006 (0.009)	0.007 (0.010)	0.005 (0.007)	Observations 3,895
R-squared (within)	0.136	0.263	0.269	0.272	0.413	
Tenant x CDD (7th Dec.)	-0.003 (0.007)	-0.004 (0.008)	0.002 (0.008)	0.002 (0.008)	0.001 (0.006)	Observations 3,871
R-squared (within)	0.108	0.209	0.227	0.231	0.327	
Tenant x CDD (6th Dec.)	0.010 (0.008)	0.019*** (0.007)	0.016** (0.006)	0.015** (0.006)	0.012*** (0.004)	Observations 3,976
R-squared (within)	0.077	0.155	0.159	0.168	0.285	
Tenant x CDD (5th Dec.)	0.003 (0.007)	0.001 (0.008)	0.000 (0.008)	-0.001 (0.008)	0.001 (0.006)	Observations 3,921
R-squared (within)	0.102	0.183	0.209	0.209	0.308	
Tenant x CDD (4th Dec.)	0.009 (0.011)	0.006 (0.012)	0.006 (0.009)	0.007 (0.009)	0.006 (0.006)	Observations 3,912
R-squared (within)	0.252	0.317	0.338	0.338	0.547	
Tenant x CDD (3rd Dec.)	-0.017 (0.014)	-0.019 (0.015)	-0.009 (0.014)	-0.008 (0.014)	-0.003 (0.008)	Observations 4,003
R-squared (within)	0.098	0.127	0.151	0.153	0.423	
Tenant x CDD (2nd Dec.)	0.006 (0.010)	0.010 (0.008)	0.013 (0.009)	0.016 (0.010)	0.012* (0.006)	Observations 3,883
R-squared (within)	0.115	0.165	0.186	0.188	0.335	
Tenant x CDD (1st Dec.)	-0.004 (0.013)	-0.006 (0.014)	-0.006 (0.012)	-0.008 (0.014)	-0.004 (0.009)	Observations 3,915
R-squared (within)	0.096	0.130	0.142	0.143	0.322	
Account & Time F.E.s, Acct. Trend	YES	YES	YES	YES	YES	
Characteristics Interactions	NO	YES	YES	YES	YES	
Characteristics Interactions w/ Year-Built	NO	NO	YES	YES	YES	
Switchers Controls	NO	NO	NO	YES	YES	

Notes: The dependent variable in columns (1)-(4) is the natural logarithm of electricity use in a billing month. The dependent variable in column (5) is the natural logarithm of billed expenditure in a billing month. Standard errors clustered at the building level are in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

costs. This is well below the payback threshold for most firms' energy conservation investments (Anderson and Newell, 2004). With a unit- or firm-level sub-meter cost of \$625, a cost which would be incurred up-front, and an average annual treatment effect of 19,000 kWh saved among high consuming firms, the cost effectiveness is 3.3 cents per kWh after the first year, 1.6 cents per kWh after two years, and 1.1 cents after 3 years, assuming the annual electricity savings persist at the same level.²⁷ The submeter costs cited above do not capture other potential costs from switching to tenant-paid contracting, such as tenants who prefer a bundled rent and utilities payment who would

27. In most states sub-meter system costs can be recovered through surcharges on tenant utility bills. This enables owners to recover their investments costs. If the owner's surcharge doesn't recover the full value of the savings, the payback period may be longer.

Figure A1: Relationship between consumption and temperature

Notes: The figure reports the coefficients of decile-by-decile regressions of normalized kWh on average daily temperature. Moving left to right, from the lower to upper deciles, there is no evidence of a monotonic relationship between consumption and temperature.

therefore pay a premium for such an arrangement. While it is difficult to test for this possibility or identify a value for these other potential costs, a full accounting of all costs should be compared to the benefits we have estimated here.

A.5 Energy Savings from Restructuring Contracts

To calculate the energy savings from restructuring rental contracts for the largest ten percent of commercial firms nationwide, we perform the following calculation. There are 130 million residential electricity customers in the U.S., of whom 10.4 million rent dwellings with utilities included (EIA, 2009). Assuming they conserve 0.7 percent of their electricity when exposed to a non-zero price (Levinson and Niemann, 2004), total residential savings are 142 million kWh per year. By comparison, there are approximately 18 million commercial sector electric customers in the U.S. (EIA, 2017), 39 percent of which rent their building space (based on the share of tenanted buildings in the U.S. in EIA, 2012). Suppose 25 percent of those (1.74 million) have an owner-paid utilities contract. The top size decile, 174,000 customers, save a total of 411 gigawatt-hours per year (1.4 percent based on our preferred empirical estimates) from a switch to tenant-paid contracts. This amounts to 289 percent of the residential sector analog. Under much more conservative assumptions, this number falls to 177 gigawatt-hours per year, or 125 percent of the residential sector analog. We reduce the fraction of renters from 39 percent to 36 percent to reflect the share of tenanted floor space, rather than the share of tenanted buildings (EIA, 2012), use the average electricity use across all large firms (not just those on owner-paid contracts, who use more electricity), and adjust our treatment effect estimate down by one standard deviation. These changes are multiplicative and thus result in an extremely conservative estimate. Importantly, this calculation does account for

underinvestment in energy efficiency, and the subsequent increase in energy consumption, that may arise from a switch from owner to tenant-paid contracts.

A.6 Monetizing External Damages

To calculate the reduction in external damages from tenant-paid contracts, we convert the estimated energy savings into avoided CO₂ and PM_{2.5} emissions, and then monetize the reduction in emissions. We do not include damages from NO_x and SO₂ emissions, given regional and federal regulations in place during our sample period. Assuming the emissions caps for these regulations were binding, a reduction in electricity consumption would not reduce aggregate emissions. While CO₂ emissions were also regulated, the early phase of this program, covering our sample period, did not have a binding cap (CRS, 2017).

To quantify CO₂ reductions we use the Environmental Protection Agency's eGRID database which provides average 2009 emission rates for the New England subregion, measured as tons emitted per MWh of electricity produced. Since we use average CO₂ emission rates in our calculations, rather than marginal rates, our estimated reductions are conservative (see Rothschild and Diem, 2009). The eGRID emission factors together with the energy savings among the top decile firms and the total number of tenant-paid firms translates into aggregate CO₂ savings of between 615 to 1200 thousand tons per year.²⁸ To give a sense of scale, this is between 3.3 to 6.6 times the average annual savings achieved from yearly Weatherization Assistance Program (WAP) retrofits.²⁹ The PM_{2.5} emission rates estimate is obtained from Connors et al. (2005). The PM_{2.5} emission rate estimate is obtained from Connors et al. (2005). Marginal CO₂ damages are from IWGSCC (2015) and damage estimates for PM_{2.5} come from Muller and Mendelsohn (2007).

28. Total CO₂ emissions saved per tenant-paid firm during summer months, using the Oster bounds, is 3.5-7 tons.

29. An average of 175,000 WAP retrofits are performed every year, which save approximately 1.06 tons of CO₂ per household per year (Fowlie et al., 2018; DOE, 2017; EIA, 2010). These retrofits therefore save 186,000 tons of CO₂ every year.



The IAEE is pleased to announce that our leading publications exhibited strong performances in the latest 2018 Impact Factors as reported by Clarivate. The Energy Journal achieved an Impact Factor of 2.456 while Economics of Energy & Environmental Policy saw an increase to 2.034.

Both publications have earned SCIMago Journal Ratings in the top quartile for Economics and Econometrics publications.

IAEE wishes to congratulate and thank all those involved including authors, editors, peer-reviewers, the editorial boards of both publications, and to you, our readers and researchers, for your invaluable contributions in making 2018 a strong year. We count on your continued support and future submission of papers to these leading publications.