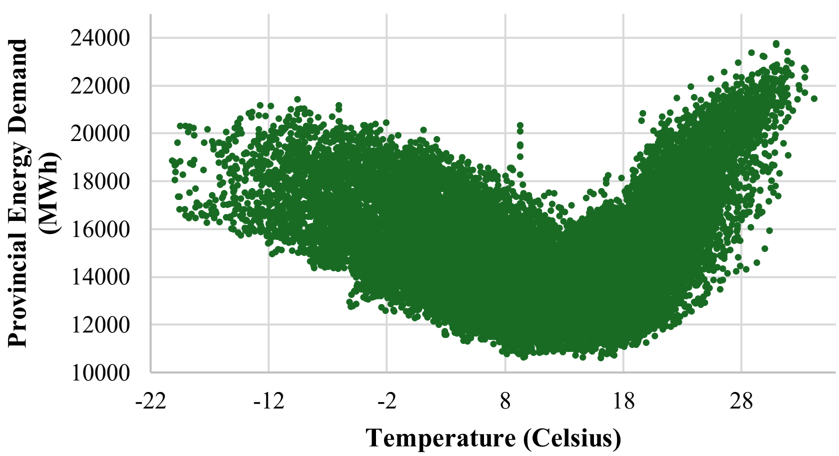
Web Appendix 1: Relationship Between Temperature and Provincial Energy Demand

This appendix documents the observed relationship between provincial energy demand and temperature, using hourly data from August 2020 to September 2022. The patterns presented here provide important context for the modeling of peak energy consumption and the design of curtailment interventions.

Figure WA-1 illustrates the intricate relationship between provincial energy consumption and temperature. Panel A displays a complex, nonlinear, and asymmetric relationship, where energy consumption increases significantly when temperatures are at the extremes (both high and low). However, consumption is more sensitive to incremental changes when the temperature is higher. Panel B presents the time series of energy consumption over several years, showing consistent patterns with peaks occurring primarily during the summer months.

Panel A: Provincial Energy Demand (MWh) vs Temperature (Celcius)



Panel B: Provincial Energy Demand (MWh) vs Time

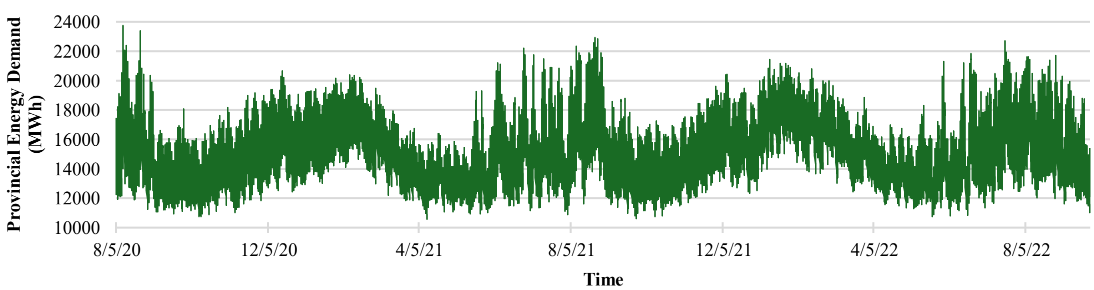


Figure WA-1: Energy Demand and Temperature Patterns

Notes: Panel A visualizes the relationship between provincial energy demand in MWh (plotted on the y-axis) and temperature in Celsius (plotted on the x-axis), using observed hourly data from August 5, 2020, to September 25, 2022. The plot displays a complex, nonlinear, and asymmetric relationship between provincial demand and temperature. Panel B plots the time series for hourly provincial energy demand in MWh from August 5, 2020, to September 25, 2022, demonstrating consistency in the consumption patterns, with the largest peaks occurring in the summer months.

Web Appendix 2: Estimation Details for Benchmark Models

This appendix describes the estimation approaches used for the benchmark forecasting models evaluated in our analysis. We detail the modeling choices and tuning procedures applied to ARIMA, regression with ARIMA errors, gradient boosting, and LSTM models. This ensures transparency and reproducibility of the comparative modeling results presented in the main paper.

ARIMA and Regression with ARIMA Errors – we estimate an ARIMA model with lagged energy demand as the independent variable using maximum likelihood estimation. When tuning the hyperparameters, we use the KPSS method to determine the optimal order of differencing (d). The remaining hyperparameters – p (order of autoregressive component), q (order of moving average component), P (order of seasonal autoregressive component), and Q (order of seasonal moving average) – are optimized by minimizing the AICc using a grid search approach based on a stepwise algorithm. We estimate the regression with ARIMA errors using the same approach. However, we include additional regressors, including temperature, temperature^2, and workday dummy, the same as in the NNTS model.

Gradient Boosting – we train a gradient boosting machine with the same predictors as those in the NNTS model: lagged energy demand, temperature, temperature^2, and workday dummy. Grid search, combined with cross validation, is used to tune the number of trees to fit, maximum depth of each tree, minimum number of observations in the trees' terminal nodes, and shrinkage rate. The final model is selected based on minimizing the out of sample error.

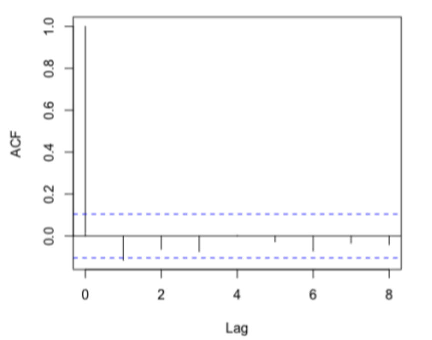
LSTM – an LSTM model is trained using lagged values of energy demand, temperature, temperature^2, and workday dummy, as predictors, for consistency with the NNTS model. The hyperparameters, including number of units in the first layer, number of units in the second layer, learning rate, and batch size, are tuned using a grid search approach, combined with cross validation. The final model is selected based on minimizing the out of sample error.

Web Appendix 3: Robustness Checks

Temporal Dependence and Repeated Treatment Bias

To test whether the treatment estimate is biased due to potentially unmodeled temporal dependence in the data, we run diagnostic tests on the residuals for up to 8 lags, which is the maximum possible lags given the nine curtailment calls. The Box-Ljung test indicates that there is no significant autocorrelation ($\chi^2(8) = 11.24$, $p = 0.1884$), which is supported by visual inspection of the Autocorrelation Function (ACF) and Partial ACF plots (see Figure WA-2 Panel A and B, respectively with dashed blues lines representing 95% CI). In combination, the lack of significant autocorrelation in the residuals and the insignificant interaction between condition and curtailment call number minimizes the concern that the treatment estimate is biased due to temporal dependencies in the data or repeated treatments, demonstrating that the results are robust.

Panel A: ACF Plot of Model Residuals



Panel B: PACF Plot of Model Residuals

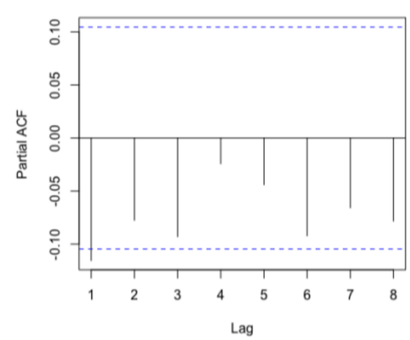


Figure WA-2: ACF and PACF Plots of Model Residuals

Comparing Curtailment Call Effectiveness 2022 vs 2021 for the Experimental Group

To further assess the robustness of the intervention, we conducted a within-subject year-over-year analysis for experimental group organizations, comparing the curtailment performance in 2021, when they received a standard email (pre-intervention), to their performance in 2022, when they received the behaviorally informed intervention email (post intervention). This temporal comparison of organization behavior between the pre- and post-intervention periods helps eliminate potential biases stemming from characteristics unique to the organizations assigned to the experimental group. This comparison becomes particularly important given that we were unable to fully randomize organizations to condition due the small sample size in the experiment.

Using full location-level data from 2021 and 2022, we estimate a linear mixed-effects model with hourly contribution to provincial energy demand as the dependent variable, dummy variables for curtailment call hours and year, along with the interaction between these variables as predictors. We include random effects for client ID for client-specific unobservable factors and control for generator ownership, hour of the day, and the month of the year. The latter is crucial since 2022 data only covers the period through August, while 2021 data spans the entire year.

Table WA-1: Year-Over-Year Impact of Curtailment Calls for the Experimental Group

		Fixed Effects	
Predictors	**Estimates**	**std. Error**	**CI (95%)**
(Intercept)	0.00579	0.00223	0.00142, 0.01016
Curtail Call Hour [Yes]	-0.00109	0.00008	-0.00124, -0.00094
Year [2022]	-0.00040	0.00001	-0.00043, -0.00038
Curtail Call Hour [Yes] * Year [2022]	-0.00038	0.00013	-0.00064, -0.00013
Generator [Yes]	0.00298	0.00378	-0.00443, 0.01039
Month		Too Many	
Hour		Too Many	
Random Effects			
$\sigma^2_{\text{Residual}}$	0.000010		
σ^2_{ID}	0.000075		
Intraclass Correlation Coefficient	0.885976		
N_{ID}	23		
Observations	334,438		

The results, presented in Table WA-1, mirror the experimental findings. Curtailment calls during the intervention period in 2022 lead to a significant reduction in energy consumption compared to the pre-intervention period in 2021 for the same locations, with an additional decrease of 0.00038 ($\beta = -0.00038, 95\% \; CI \; [-0.00064, -0.00013]$). This represents a 34.86% (95% CI [11.93%, 58.72%]) reduction in energy demand, building on the baseline curtailment effect of 0.00109, closely in line with the reduction between control and experimental groups during the intervention. Collectively, these findings provide further evidence of the effectiveness of the behaviorally informed emails in reducing energy consumption.

Limited Sample Size

Given the limited sample size (i.e., 9 curtail calls across 39 clients), we assess the robustness of our results using three additional complementary approaches. First, we employ a semi-parametric residual *bootstrapping* approach for linear mixed effects models (Carpenter et al. 2003), which allows for flexible estimation of CIs in small-sample, hierarchical data. The bootstrapped results indicate a significant treatment effect ($\beta\ =\ 259.33, 95\%\ CI[-477.48, -54.78]$) with tighter confidence bounds than those in the main analysis. We also obtain consistent results using parametric, wild, and random effects block (REB) bootstrapping, demonstrating the robustness of the results. Second, we conduct a non-parametric permutation test suitable for small samples (Good 2005). We generate a null distribution of treatment effects by randomly reassigning treatment labels across clients 1,000 times, re-estimating the effect for each permutation. Comparing the observed effect to this null distribution reveals that only 1.5% of permuted datasets produce an effect as extreme or more extreme than the observed effect ($\beta\ =\ 259.33, p = 0.015$), indicating that the result is highly unlikely to occur by chance. Third, we perform leave-one-out sensitivity analysis (Cook 1977) to evaluate whether the treatment effect is driven by a small number of influential clients. We iteratively remove each client from the sample, refit the mixed effects model, and compare the resulting estimates to the full-sample effect. Across subsamples, the treatment effect remains robust ($M\ =\ 254.86, SD = 45.92, 95\%\ CI[-325.91, -103.61]$), with the sign and magnitude consistently supporting the main finding. These three complementary robustness checks mitigate concerns about the limited sample size and strengthen confidence in the validity of our main results.

References:

Carpenter JR, Goldstein H, Rasbash J (2003) A novel bootstrap procedure for assessing the relationship between class size and achievement. Journal of the Royal Statistical Society: Series C (Applied Statistics) 52(4):431–443.

Cook RD (1977) Detection of Influential Observation in Linear Regression. Technometrics 19(1):15–18.

Good P (2005) Permutation, Parametric, and Bootstrap Tests of Hypotheses (Springer, New York, NY).

Web Appendix 4: Multi-phase follow-up experiment

4.1 Procedure

To better understand which features of the emails may have driven the differences observed between the control and experimental messages in Study 2, we conducted a pre-registered, incentive-compatible, multi-phase follow-up experiment (https://aspredicted.org/q8qq-p498.pdf). We recruited 330 US participants with 5+ years of management experience from Prolific Academic to participate in a two-part study on energy pricing communications, which included a training and testing phase. A total of 291 participants (44.7% female, M_{age} = 46.61) completed both parts and were included in the analyses.

Part 1, our training phase, aimed to inform participants about the context, critical peak pricing, and introduce the strategies that managers can use to reduce costs under such a pricing structure. After providing consent, all participants were first asked to imagine that they were managers of a large organization who deals with a critical peak pricing energy scheme (a type of demand-based pricing). They were told they would be trained (and tested) on how the program works and given advice on how to reduce energy consumption to avoid the high costs during peak hours. They were also told that they could earn bonus payments based on their responses.

Participants were then provided with detailed information about critical peak energy pricing and how it works (see Table WA-2 for the information provided). To ensure that all participants understood the context and pricing scheme, they were then asked four knowledge questions based on the content. Participants could only proceed after answering each question correctly. If a question was answered incorrectly, the relevant information was shown again, and they were allowed to retry. The majority of participants (74.9%) answered all questions correctly on their first attempt. Fewer than 4% required more than three attempts to answer a question correctly.

Next, participants were asked to imagine they had hired a consulting firm to help them manage their company's energy costs. We then aimed to simulate the type of training that the companies in our main study would have received from En-Pro (i.e., our partner organization). Participants were reminded that in approximately a week's time, during Part 2 of the study, they would have the opportunity to earn bonus payments for correctly recalling and answering questions related to this training. This incentive compatible design mirrors the real-world financial rewards managers may receive when making similar cost-saving decisions.

Participants were then presented with training information that 1) explained the impact of peak energy consumption on company costs, 2) detailed how the consulting firm would send advanced email notifications before predicted peak periods, and 3) outlined recommended actions to reduce energy use during those times (see Table WA-3 for the complete information). Finally, participants reported demographic and background information including their age, gender, years of management experience, whether they had previously been responsible for energy-related decisions at work, and their familiarity with demand-based pricing programs. A total of 83.8% of the sample reported more than 5 years of management experience, and 64.4% had been involved in energy-usage-related decisions in the workplace. On average, participants reported low to moderate familiarity with demand-based pricing programs (M = 2.694, SD = 1.060 on a 5-point scale).

One week later, participants were invited to complete Part 2, the testing phase of the study. The purpose of Part 2 was to empirically examine how the control and experimental emails affected participants' understanding of strategic energy decisions they should make (with bonuses awarded for correct responses), their between-subjects perceptions of the emails, and their thoughts on which components of the emails would be most effective at motivating organizations to curtail energy use.

Specifically, in Part 2, participants were again asked to imagine they were the managers of the large company subject to critical peak pricing, who had hired the consulting firm to help minimize their energy costs. They were then asked to imagine that they received an email from the consulting firm. Participants were then randomly assigned to receive either the control email or the experimental email (see Figure WA-3).

After reading the email, participants were then asked to: 1) identify the correct times to curtail energy consumption (i.e., when to act), and 2) identify the appropriate energy saving actions they could take based on the training provided in Part 1 and the information in the email (i.e., how to act). Participants were informed they would receive a $0.10 bonus for each correct response.

Next, participants evaluated the email message on several dimensions including: 1) clarity (e.g., how clearly does this email explain that a curtail has been called today?, 1 – Not at all clearly; 7 – Very clearly), 2) effectiveness (e.g., how effective do you believe the current email message would be in driving energy-saving curtailment behaviors?, 1 – Not at all effective; 7 – Very effective) and 3) likelihood of manager action (e.g., what is the likelihood that a manager in this context would take effective actions to reduce their energy consumption within the given time window upon receiving this email?, 1 – Not at all likely; 7 – Very likely).

Participants were then asked, using a hot spot tool, to select which components of the email (e.g., subject line, visual elements, individual sentences) they believed were most impactful for encouraging firms to effectively reduce peak energy consumption. They were first asked to select (up to) three components that were most impactful, then were asked to select the one element that was most impactful. Participants were thanked and informed of their bonus payments.

4.2 Results

First, to assess whether the experimental email (vs. control) improved participants' understanding of when and how to act, we compared the proportion of correct responses to the curtailment timing and energy-saving actions questions. Across conditions, we observed no difference in participants' ability correctly identify the curtailment window (95.2% in the control condition vs. 92.4% in the experimental condition; $\chi^2(1) = 1.038$, $p = .308$). However, participants in the experimental condition were significantly more likely to correctly identify all three recommended energy-saving actions (70.8%) compared to those in the control condition (51.7%; $\chi^2(1) = 11.210$, $p < .001$). More specifically, participants were equally likely to identify turning off / down the air conditioning as one of the correct actions to take ($\chi^2(1) = 2.521$, $p = .112$). However, participants in the experimental condition were significantly more likely to recognize reducing non-essential operations ($\chi^2(1) = 5.122$, $p = .024$) and using a back-up power generator ($\chi^2(1) = 12.129$, $p < .001$) as appropriate strategies during curtailment hours. These findings suggest that while both emails effectively communicated *when* to act, the experimental email may have improved participants understanding of how to act by clarifying which specific actions should be taken.

Next, to examine whether participants' perceptions of the email messages differed by condition, we conducted a series of one-way ANOVAs on our perception measures. Compared to the control condition, participants rated the experimental email as significantly clearer ($M_{experimental} = 6.465$, $SD = .747$; $M_{control} = 6.020$, $SD = 1.037$, $F(1, 289) = 17.572$, $p < .001$, $\eta_p^2 = .057$), more effective ($M_{experimental} = 6.042$, $SD = .900$; $M_{control} = 5.449$, $SD = 1.195$; $F(1, 289) = 22.784$, $p < .001$, $\eta_p^2 = .073$), and more likely to encourage managers to take action ($M_{experimental} = 6.000$, $SD = .931$; $M_{control} = 5.449$, $SD = 1.183$, $F(1, 289) = 19.438$, $p < .001$, $\eta_p^2 = .063$).

Finally, we explored which components of each email were perceived as most impactful in motivating firms to effectively reduce peak energy demand. Figure WA-4 presents the regions in each letter that were selected by more people than would be predicted by chance.[1] As shown, the curtailment time window was considered especially important in both emails—selected 82.3% of the time in the control email ($\chi^2(1) = 324.265$, $p < .001$) and 51.4% of the time in the experimental email ($\chi^2(1) = 88.674$, $p < .001$). In fact, the time window was selected as the most important component of both the control (61.2%) and experimental (41.7%) emails.

In the control email, other elements selected more often than predicted by chance included the sentence specifying the day's predicted energy demand (50.3%, $\chi^2(1) = 73.195$, $p < .001$) and the directive "We strongly recommend acting for today's entire window if you can" (40.0%, $\chi^2(1) = 9.769$, $p = .002$).

On the other hand, in the experimental email, each of the three planning prompt actions (i.e., *how to act*) were considered effective (turning down air-conditioning: 50.0%, $\chi^2(1) = 81.000$, $p < .001$; reducing non-essential operations: 45.1%, $\chi^2(1) = 56.877$, $p < .001$; and using a back-up generator:

[1] Given that participants were asked to choose up to three components out of the 14 in the control condition and 15 in the experimental condition, this means looking into elements chosen more often than 21.4% of the time in the control and 20% of the time in the experimental condition.

32.6%, $\chi^2(1) = 14.377$, $p < .001$). The meter visual (with an arrow pointing in the red zone), was also rated as impactful, albeit to a lesser extent (30.6%, $\chi^2(1) = 10.028$, $p = .002$).

Interestingly, in both emails the subject line was selected less frequently than would be expected by random chance (control email: 2.7%, $\chi^2(1) = 30.492$, $p < .001$; experimental email: 6.3%, $\chi^2(1) = 17.016$, $p < .001$).

4.3 Discussion

Overall, this multi-phase, incentive-compatible experiment presents two key insights. First, for organizations subject to CPP, knowing *when* to act appears to be the most crucial driver of effective behavior. In this follow-up experiment, participants across both conditions accurately identified the curtailment window, and this did not differ by condition. Moreover, the curtailment time window was selected as the single most important element of the email across both conditions. These results, together with those from our main study—where the control email consistently supported curtailment—suggest that clearly communicating *when* to curtail at a key point in the decision-making process, may be a foundational element for driving action.

Second, these findings suggest that the primary advantage of the experimental email was its ability to reinforce the specific actions managers can take (how to act), through the use of planning prompts. Participants who received the experimental email (vs. control) were significantly more likely to correctly identify the recommended energy-saving strategies. Moreover, when asked to indicate the most impactful elements of the email, participants in the experimental condition were especially likely to select the three planning prompt actions as critical. Consistent with this, the experimental email was rated more positively in terms of clarity, perceived effectiveness and likelihood of prompting manager action. Taken together, these findings suggest that the greater curtailment observed in the experimental condition of our field study may have been driven, at least in part, by the inclusion of specific planning prompts that made the appropriate actions more salient and easier to implement at the point of decision-making. Even when managers are trained in advance, it is possible—just as we observed in this follow-up study—that relevant actions are not always top-of-mind when a curtailment call arrives. As such, planning prompts may serve as effective cognitive aids that facilitate behavior.

While these findings provide inital insight into why the experimental emails may have been particularly effective, it is important to acknowledge that this study was hypothetical and conducted with a different participant sample. Ideally, we would have been able to experimentally isolate these elements in the field to determine which components were necessary for effective behavior change. Nevertheless, the multi-phase design of the study simulated key aspects of the main field experiment—beginning with a training phase and followed by curtailment calls. Moreover, the study recruited participants with at least 5 years of management experience, many of whom had previously been responsible for managing their organizations' energy consumption[2]. Additionally, the use of financial incentives for correctly identifying curtailment windows and energy-saving actions paralleled the real-world stakes that managers may face in reducing energy costs. While this follow-up study is not a perfect analogue to real-world operations, the results provide useful insight into the likely drivers of our interventions' effects. Future field work is needed to isolate and confirm the underlying causal mechanisms under operational conditions.

Table WA-2 Introduction to critical peak pricing in Phase 1

General information on critical peak pricing	Critical peak pricing (CPP) is a system where electricity costs more during the busiest times of the year when demand is very high. The goal of CPP pricing is to shift energy demand away from the busy times, to times when demand is lower, so that there is enough supply overall and the energy supplied can come from more sustainable sources. In this context, under critical peak pricing, large organizations' energy costs are determined based on two factors: 1) the regular (wholesale) cost of energy 2) a global adjustment fee (an additional fee / penalty that companies pay based on how much energy they were using during peaks in energy demand the previous year) These energy "demand peaks" are determined retroactively, by looking at how much energy was being consumed at the grid level (e.g., across an entire state) - <u>every hour</u> in a given year. Once the year is over, the

[2] Our results did not differ when looking only at managers with energy management experience.

	energy company identifies <u>the 5 hours during the year when the grid level energy consumption was highest</u>, and those 5 hours are labeled as the demand peaks. To determine an organization's global adjustment fee, the energy company then looks at how much energy each organization was using during the demand peaks (% of energy they were consuming out of total grid consumption). They then multiply the % of total peak energy an organization used by the global adjustment cost.
Example	For example, if Company X's energy usage during the peak hours of the year was 100 kWh, and that accounted for 5% of the total energy demand in the region during those peak hours, they would pay 5% of the global adjustment cost. If the global adjustment costs were $10 Million, the organization would pay <u>both</u> 1) their wholesale price for <u>all</u> their electricity used, as well as 2) their global adjustment fee of $500,000 (i.e., 5% of 10 Million). Therefore, if a company is a large contributor to demand peaks, they will pay a higher global adjustment fee on top of what they are already paying for energy.

Table WA-3 Training information from consultant in Phase 1

General information	The first thing you learn from the consultant is that over the past several years, for large organizations such as yours, the Global Adjustment cost has become the largest portion of the electricity bill and currently accounts for approximately 70% to 80% of the total electricity bill! Therefore, the consultant advises that focusing on **reducing your company's energy consumption during periods of peak demand** is the *most efficient* way to effectively reduce energy costs. However, periods of peak demand are only determined at the end of the year after the peaks have already happened, making it difficult to know when to reduce energy consumption. To solve this problem the consultancy offers an advisory service where they monitor and track energy demand at the grid level (along with weather and other important factors). They do so in order to anticipate and forecast potential peaks in energy demand. By forecasting and communicating these potential demand peaks, your company can have advanced warning of when peak demand hours might occur, giving you the ability to reduce your energy consumption during those times. The consultancy uses a sophisticated artificial intelligence algorithm to predict when peaks in energy demand are likely to occur. Each time a potential demand peak is predicted, they will **<u>provide your company with an email message</u>**, advising you to **reduce your energy consumption** (i.e., **curtail your consumption**) during the predicted time

	frame. Shifting energy demand even modest amounts during peak time can lead to substantial energy savings. For example, reducing demand by a small amount during two peaks hours for a large organization like yours could save the organization over $100,000. It is important to note that it is difficult to predict (in advance) with certainty which 5 hours will ultimately be demand peaks, and therefore, on any given curtailment day the consultancy provides **windows of time that they recommend you curtail your energy consumption**, to increase the chances you reduce your energy during peak hours. It is advised that, if you can, you curtail your energy during the entire time frame.
Specific actions (after page break)	Of course, it is crucial for you to know what to do when such a curtailment email arrives. So in addition to providing you emails recommending when to act, the consultancy also provide advice on some of the **<u>most effective actions to take</u>** to reduce peak energy demand. See below for the recommended actions: **How to Cut Electricity Use During Peak Demand Hours** Peak electricity demand occurs on the hottest summer days, when air conditioners and heavy machinery strain the power grid. Reducing usage during these times can lead to substantial cost savings while easing pressure on the grid. Based on our review of your operations, here are the most effective steps you can take: **1. Reduce Air Conditioning** Your cooling systems are one of the biggest energy draws during peak hours. To cut usage without sacrificing comfort: • Raise the thermostat **by 2-3 degrees** during peak hours. • **Pre-cool the building earlier in the day** so that cooling systems work less when demand (and pricing) is highest. • Use fans, blinds, and weatherproofing to maintain comfortable temperatures with less energy. **Why?** Air conditioning makes up a large portion of your energy costs during summer peaks. Adjusting how and when you cool can lead to big savings while still keeping employees comfortable. **2. Shift or Reduce Energy-Intensive Operations** We recommend adjusting work schedules so high-energy tasks are completed **before or after peak hours**. This could mean: • Reducing or delaying the use of **heavy machinery** and non-essential equipment. • Shifting certain operations to **off-peak times** to even out energy use.

Why? Powering down large equipment during peak times avoids unnecessary energy costs and lowers the overall electricity bill.

3. Use On-Site Backup Power

If you have a **generator or battery system**, using it during peak hours can dramatically reduce your electricity costs.

Why? Drawing less power from the grid at peak times means lower demand charges and protection from potential supply issues.

4. Identify Where You are Using Energy Most
Right now, it's unclear where all your highest energy loads are coming from. We strongly recommend taking steps to target energy waste so you can cut consumption without disrupting operations. Some ways to do this include:

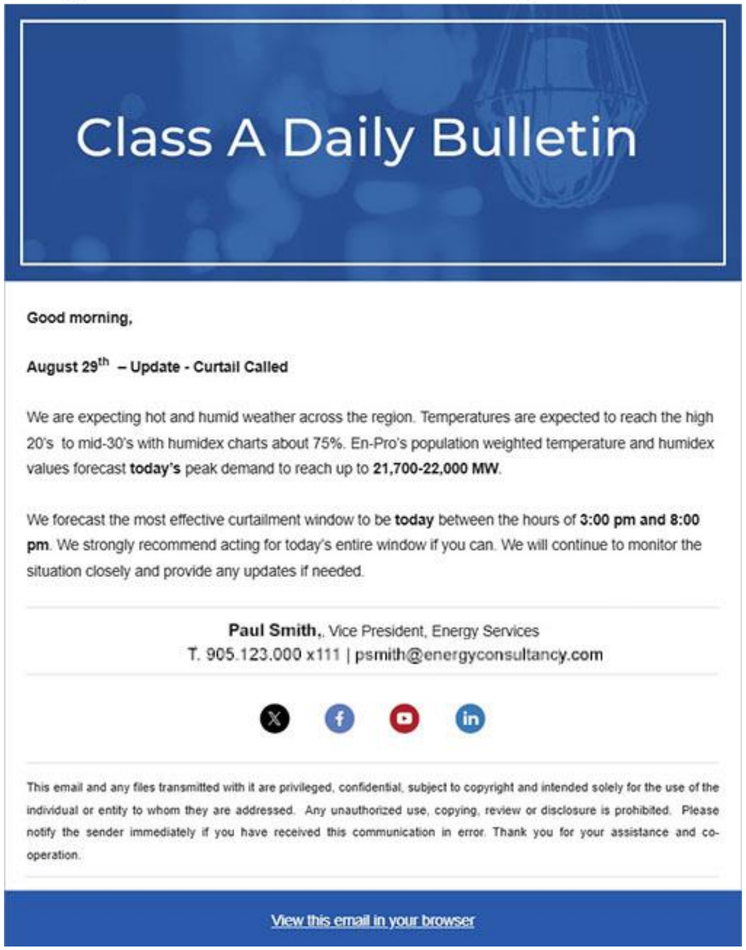
- Identifying non-essential equipment and **powering it down during peak hours**, such as lighting, office equipment, or idle machines.
- **Using circuit meter technology** to pinpoint where energy is being used most.
- Looking upstream in your manufacturing process—**reducing energy use in areas that won't slow down production elsewhere**.

Why? Companies that take the time to identify where they're using the most energy tend to see the biggest savings. If you take this step, you'll have clear data to drive smarter decisions and maximize your energy savings.

By acting on these recommendations, your company can significantly **lower electricity costs, improve energy efficiency, and reduce strain on the grid**—all while maintaining operations.

Control email

Subject line: Class A Daily Bulletin - August 29, 2022



Experimental email

Subject line: Curtail called for today August 29th 3-8pm

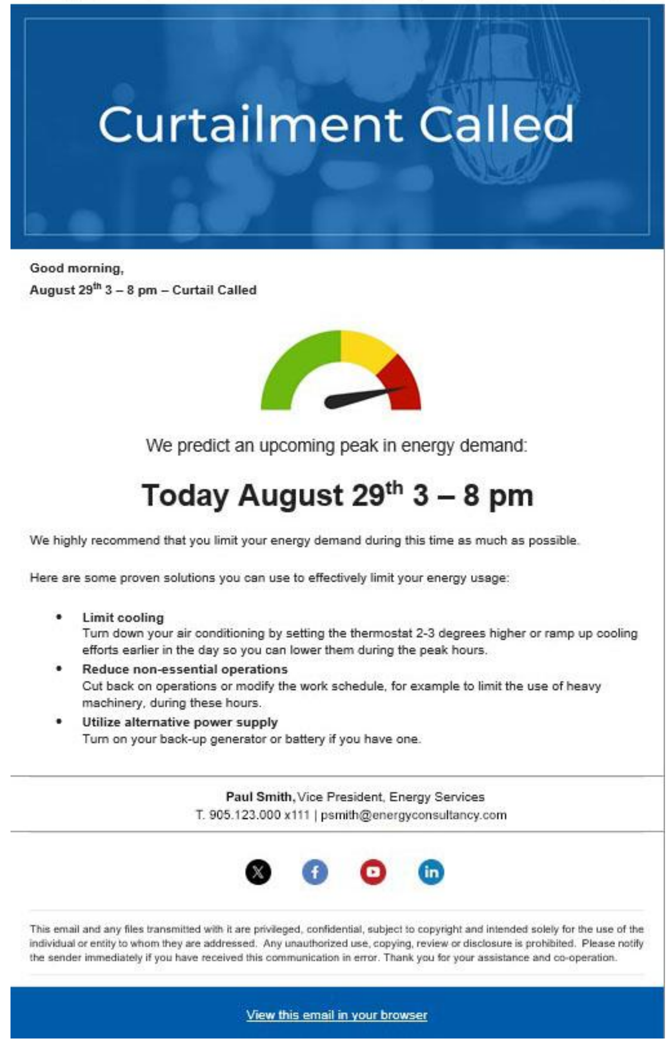
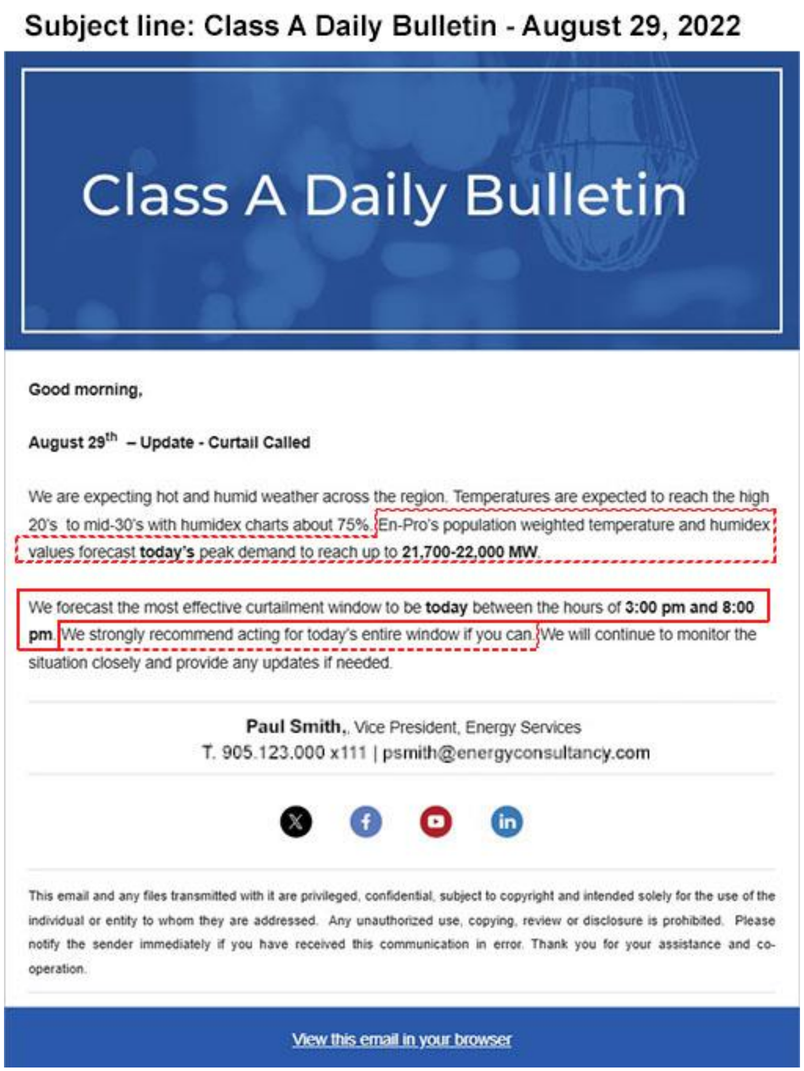


Figure WA-3 Emails for control and experimental condition

Control email



Experimental email

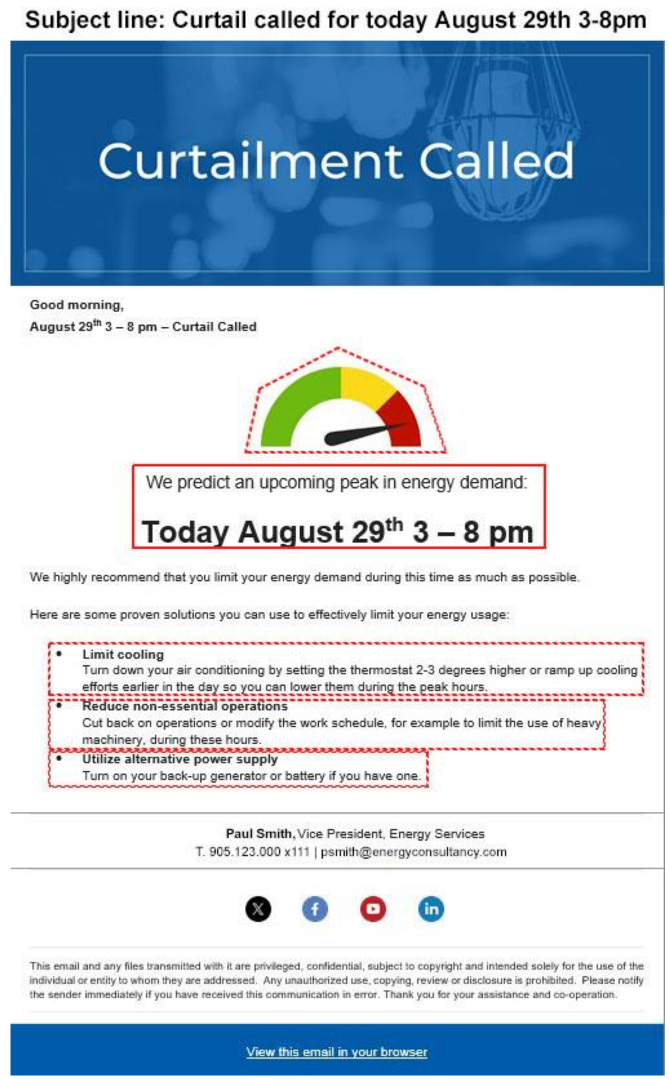


Figure WA-4 Most selected areas by experimental condition

Web Appendix 5: Assumptions and Calculations for Implementation Cost Estimates

This appendix documents the detailed assumptions, data sources and step-by-step calculations used to estimate the annual implementation costs of our intervention for clients with and without generators. These costs reflect the two primary pathways through which clients adjust energy consumption during curtailment events: (1) by substituting grid electricity with on-site generation, and (2) by opportunity costs from curtailing profit-generating activities. We also quantify the proportion of these costs attributable to false alarms, enabling full transparency and reproducibility of our net savings estimates.

We base our calculations on the actual characteristics of curtailment hours in 2022, which included 38 curtailment hours (Table 4, Panel B), 16 of which were false alarms (42%).

<u>Step 1. Generator-based curtailment costs (for clients with generators)</u>

1. Required generator capacity is 693 kW, calculated as:

- Advanced AI-driven curtailment: 427.20 kWh per hour (Table 2 for curtailment values and Table 4 Panel B for average peak demand over the top five peaks)
  (0.00121% + 0.00073%) × 22,012,000 kWh/hour average provincial peak demand)
- Behavioral Nudging-driven curtailment: 265.54 kWh per hour (Table 6)
- Total required generator capacity: 693 kW

2. Capital costs per kW of installed capacity (U.S. Department of Energy's National Renewable Energy Laboratory 2019, 2023):

- Diesel: USD $800/kW
- Natural gas: USD $1,000/kW
- Battery: USD $1,928/kW

   Assuming a 30-year generator lifespan and an exchange rate of CAD $1.39/USD, annual capital costs are:

- Diesel: CAD $25,659/year
- Natural gas: CAD $32,074/year
- Battery: CAD $61,838/year

3. Maintenance costs per kW (U.S. Department of Energy's National Renewable Energy Laboratory 2019; Electric Power Research Institute 2020):

- Diesel and natural gas: USD $405/kW
- Battery: USD $180/kW

Converted to CAD and annualized over 30 years:

- Diesel: CAD $12,990/year

- Natural gas: CAD $12,990/year
- Battery: CAD $5,773/year

4. Raw material (fuel) costs (U.S. Department of Energy's National Renewable Energy Laboratory 2019):

- Diesel: USD $0.200/kWh
- Natural gas: USD $0.124/kWh
- Battery: USD $0.00/kWh

Multiplying by 693 kWh/hour × 38 curtailment hours in 2022 (as per Table 4, Panel B), and converting to CAD:

- Diesel: CAD $7,313/year
- Natural gas: CAD $4,534/year
- Battery: CAD $0/year

5. Total generator costs (capital + maintenance + raw materials):

- Diesel: CAD $45,962/year
- Natural gas: CAD $49,598/year
- Battery: CAD $67,611/year

6. Weighted average generator cost:
Assuming an even distribution of generator types (1/3 diesel, 1/3 natural gas, 1/3 battery):

- (1/3 × $45,962) + (1/3 × $49,598) + (1/3 × $67,611) = CAD $54,390/year per client with generators

<u>Step 2. Opportunity costs (for clients without generators)</u>

1. Profit per kWh curtailed:

- Gross operating surplus + gross mixed income in Ontario = CAD $390.19B (Statistics Canada 2024)
- Annual energy demand from organizations in Ontario = 90.80B kWh (Faruqui et al. 2007; IESO 2023)
- Profit per kWh curtailed = $390.19B ÷ 90.80B kWh = CAD $4.30/kWh

2. Hourly curtailment without generators:

- Advanced AI-driven curtailment: 266.50 kWh per hour (Table 2 for curtailment values and Table 4 Panel B for average peak demand over the top five peaks)
(0.00121% × 22,012,000 kWh/hour average provincial peak demand)
- Behavioral Nudging-driven curtailment: 265.54 kWh per hour (Table 6)
- Total curtailed: 532 kWh per hour

3. Proportion of curtailment from profit-generating activity:

- Assumed to be 50 percent.

4. Opportunity cost calculation:

- \$4.30/kWh × 38 curtailment hours (as per Table 4, Panel B) × 532 kWh/hour × 50% = CAD \$43,442/year per client without generators

Step 3. Impact of false alarms

1. In 2022, there were 38 curtailment hours, of which 16 were on days with false alarms (Table 4, Panel B).

- Proportion of false alarms = 16 ÷ 38 = 42%.

2. For clients without generators, 42 percent of the opportunity cost is attributable to false alarms:

- 42% of CAD \$43,442 = CAD \$18,246/year.

3. For clients with generators, 42 percent of the raw material cost is attributable to false alarms:

- Diesel: 42% × \$7,313 = \$3,071
- Natural gas: 42% × \$4,534 = \$1,904
- Battery: 42% × \$0 = \$0
- Weighted average false alarm fuel cost across generator types is similarly calculated.

Step 4. Summary of annual implementation costs and net savings

1. Average annual cost of implementation:

- Clients without generators: CAD \$43,442/year
- Clients with generators: CAD \$54,390/year

2. Net savings (after implementation costs):

- Clients without generators: CAD \$209,604/year per client
- Clients with generators: CAD \$274,342/year per client
- Weighted average across all clients: CAD \$232,843/year per client

3. Observations:

- False alarms significantly affect costs for clients without generators, due to the variable nature of opportunity costs.
- For clients with generators, fixed capital and maintenance costs dominate total costs, while raw material costs (and their sensitivity to false alarms) are smaller in relative terms.

References:

Electric Power Research Institute (2020) Battery Energy Storage Lifecycle Cost Assessment Summary. Report. https://www.epri.com/research/products/000000003002020048.

Faruqui A, Hledik R, Newell S, Pfeifenberger J (2007) The power of five percent: how dynamic pricing can save $35 billion in electricity costs. Report, The Brattle Group Inc., Cambridge, MA, USA. https://www.brattle.com/insights-events/publications/the-power-of-five-percent-how-dynamic-pricing-can-save-35-billion-in-electricity-costs/.

IESO (2023) Hourly Demand Report. Report. https://reports-public.ieso.ca/public/Demand/.

Statistics Canada (2024) Table 36-10-0221-01 Gross domestic product, income-based, provincial and territorial, annual (x 1,000,000). Report. https://doi.org/10.25318/3610022101-eng.

U.S. Department of Energy's National Renewable Energy Laboratory (2019) A Comparison of Fuel Choice for Backup Generators. Report. https://www.nrel.gov/docs/fy19osti/72509.pdf.

--- (2023) Cost Projections for Utility-Scale Battery Storage. Report. https://www.nrel.gov/docs/fy23osti/85332.pdf.

Web Appendix 6: Immediate Reductions in Carbon Footprint Assumptions and Calculations

This appendix documents the assumptions and calculations for clients' energy consumption in Phases 1-3 (i.e., baseline, advanced AI, advanced AI & behavioral nudging) reported in Section 4.2.1 of the paper, to ensure reproducibility of the findings and transparency of the results.

Step 1. Calculate baseline energy consumption for clients (Phase 1)

- Baseline energy consumption for clients without generators: 1,422.61 kWh per hour (Table 2 for contribution values and Table 4 Panel B for average peak demand over the top five peaks) (0.00646% contribution to provincial demand × 22,012,000 kWh/hour average provincial peak demand)
- Baseline energy consumption for clients with generators: 1,358.84 kWh per hour (Table 2 for contribution values and Table 4 Panel B for average peak demand over the top five peaks) ((0.00646%% - 0.00029%) contribution to provincial demand × 22,012,000 kWh/hour average provincial peak demand)
- Baseline energy consumption for clients (weighted average across those without (25) and with (14) generators: 1,399.72 kWh per hour
  (25/39 x 1,422.61 kWh per hour) + (14/39 x 1,358.84 kWh per hour)

Step 2. Calculate energy consumption with advanced AI (Phase 2)

- Energy consumption with advanced AI for clients without generators: 1,156.11 kWh per hour (Table 2 for curtailment values and Table 4 Panel B for average peak demand over the top five peaks)
  (1,422.61 kWh per hour baseline energy consumption - (0.00121% reduction in contribution due to advanced AI × 22,012,000 kWh/hour average provincial peak demand))
- Energy consumption with advanced AI for clients with generators: 931.63 kWh per hour (Table 2 for curtailment values and Table 4 Panel B for average peak demand over the top five peaks)
  (1,358.84 kWh per hour baseline energy consumption - ((0.00121% + 0.00073%) reduction in contribution due to advanced AI × 22,012,000 kWh/hour average provincial peak demand))
- Energy consumption with advanced AI for clients (weighted average across those without (25) and with (14) generators: 1,075.53 kWh per hour
  (25/39 x 1,156.11 kWh per hour) + (14/39 x 931.63 kWh per hour)

Step 3. Calculate energy consumption with advanced AI and behavioral nudging (Phase 3)

- Energy consumption with advanced AI and behavioral nudging for clients without generators: 890.57 kWh per hour (Table 6 and Figure 5)
  (1,156.11 kWh per hour energy consumption with advanced AI - 265.54 reduction in energy consumption due to behavioral nudging)
- Energy consumption with advanced AI and behavioral nudging for clients with generators: 666.10 kWh per hour (Table 6 and Figure 5)
  (931.63 kWh per hour energy consumption with advanced AI - 265.54 reduction in energy consumption due to behavioral nudging)
- Energy consumption with advanced AI and behavioral nudging for clients (weighted average across those without (25) and with (14) generators: 809.99 kWh per hour
  (25/39 x 890.57 kWh per hour) + (14/39 x 666.10 kWh per hour)

Web Appendix 7: Environmental Impact Assumptions and Calculations

This appendix documents the assumptions and calculation steps used to derive the environmental impact estimates reported in Section 4.2. The main paper presents the resulting emissions reductions; here we provide the underlying data sources and computational details to ensure transparency and reproducibility.

The analysis proceeds in four steps: (1) estimate organizational peak demand, (2) apply the observed reduction rate, (3) calculate avoided emissions, and (4) account for generator-related emissions and net out results.

Step 1. Estimate organizational peak demand

- Average organizational energy demand during the five annual system peaks is 14,528 MWh for Ontario and 55,877 MWh for Canada (National Resources Canada 2024; IESO 2023; Faruqui et al. 2007).

Step 2. Apply observed reduction rate

- The observed reduction rate from the intervention is 42.13 percent (Figure 8)
- Applying this reduction rate yields projected demand reductions of:
  - 6,121 MWh in Ontario
  - 23,542 MWh in Canada

Step 3. Calculate avoided grid emissions

- The marginal emissions factor for natural gas generation during peaks is 450 gCO_2e per kWh (Ontario Power Generation 2020).
- Multiplying the demand reductions by this factor yields estimated hourly avoided emissions of:
  - 2,754 tonnes of CO_2 in Ontario
  - 10,594 tonnes of CO_2 in Canada

Step 4. Account for generator-related emissions and compute net CO_2 reduction

- Estimated 34.85 percent of organizational demand is from firms with generator capacity[3]
- The observed curtailment rate for firms with generators is 50.98 percent (Figure 8).
- This yields generator-based curtailment of:
  - 2,581 MWh in Ontario
  - 9,927 MWh in Canada
- Generator emissions factors are:
  - Diesel: 800 gCO_2e per kWh (United Nations 2022)
  - Natural gas: 450 gCO_2e per kWh (Ontario Power Generation 2020)
  - Battery: 0 gCO_2e per kWh (United Nations 2022)
- Assuming an even distribution across generator types, the average generator emissions factor is 417 gCO_2e per kWh.

[3] Under the assumption that our sample of clients is representative of organizations across the province and country with respect to generator ownership

- Multiplying this factor by generator-based curtailment yields additional generator-related emissions of:
    - 1,075 tonnes of CO_2 per hour in Ontario
    - 4,136 tonnes of CO_2 per hour in Canada
- Net hourly CO_2 reductions (avoided grid emissions minus generator-related emissions):
    - 1,679 tonnes in Ontario
    - 6,458 tonnes in Canada

Household energy equivalence calculation

- The average Ontario household consumes approximately 23.28 kWh per day (Financial Accountability Office of Ontario 2019).
- The projected demand reduction of 23,542 MWh in Canada, which in the best-case scenario (when all organizations with generators are using battery generators) results in a 10,594 tonne net reduction in CO_2 emissions per hour, is equivalent to approximately 1,011,392 households' daily energy use.

These calculations provide transparency regarding the assumptions underlying our environmental impact estimates and support the robustness of the main results presented in the paper.

References:

Faruqui A, Hledik R, Newell S, Pfeifenberger J (2007) The power of five percent: how dynamic pricing can save $35 billion in electricity costs. Report, The Brattle Group Inc., Cambridge, MA, USA. https://www.brattle.com/insights-events/publications/the-power-of-five-percent-how-dynamic-pricing-can-save-35-billion-in-electricity-costs/.

Financial Accountability Office of Ontario (2019). Home Energy Spending in Ontario. Toronto, Canada. https://www.fao-on.org/web/default/files/publications/FA1911%20Home%20Energy/Home%20Energy%20Spending%20in%20Ontario%202019%20Update.pdf.

IESO (2023) Hourly Demand Report. Report. https://reports-public.ieso.ca/public/Demand/.

Natural Resources Canada (2024) Energy Fact Book. Report. https://energyinformation.canada.ca/sites/default/files/2023-10/energy-factbook-2023-2024.pdf.

Ontario Power Generation (2020). Building a brighter tomorrow. Report. https://www.opg.com/documents/opg-climate-change-plan-2020/.

United Nations (2022). Methodological tool Default values for common parameters. Report. https://cdm.unfccc.int/methodologies/PAmethodologies/tools/am-tool-33-v2.0.pdf.