Measuring Gas Demand Response Program Impacts: A Pilot Evaluation Using Thermostat Telemetry Data

Ari Kornelis, Cadmus Group, Portland, OR
Emily Kremser, Cadmus Group, Portland, OR

ABSTRACT

In recent years, extreme cold weather events have strained natural gas distribution systems across the Midwest, Northeast, and in Texas. In several cases, these weather conditions have contributed to or coincided with equipment failures leading to catastrophic service failures. These incidents have raised questions about the potential for gas demand flexibility and demand response (DR) programs to improve natural gas distribution resilience and reliability.

During winter 2020/21, a large electric and natural gas distribution utility in the Midwest conducted a gas DR pilot program involving smart thermostat direct control treatments for residential and small commercial customers. This paper presents the results of the pilot impact evaluation.

The limited temporal resolution of typical gas metering infrastructure raises challenges for measuring the impact of 2–4-hour DR events. To address this challenge, hourly consumption impacts were measured in heating system runtime minutes using thermostat telemetry data. We present a methodology for translating these heating minute impact estimates to units of gas consumption by estimating the effect of a heating system runtime minute on cubic feet of gas consumed in a second stage model.

The study provides several new insights. It demonstrates the potential for gas demand response in a cold climate Midwest region, the efficacy of experimental and non-experimental methods for gas DR impact measurement, and it provides a solution for measuring gas DR program impacts with limited advanced gas metering infrastructure.

Introduction

Gas distribution utilities might consider conducting natural gas demand response for a variety of reasons. The most common are likely to defer infrastructure investment and increase reliability and resilience of the gas distribution system. Deferred infrastructure investment may reduce system costs and provide indirect climate and environmental benefits through avoided carbon lock-in (Sato, 2021).

Theoretically, gas demand response that delivers reductions in gas consumption during peak hours can reduce the capacity required to serve customers on a distribution node. There are few studies that provide estimates of the capacity value of gas demand response. One study of Central Hudson's service territory used a probabilistic demand forecasting approach to quantify the likelihood that gas demand response could help avoid a loss of pressure incident that could trigger a distribution capacity upgrade. The study found that on the most constrained distribution node, the capacity value from gas DR could be as great as $1,000/Ccf-year (Bode, 2020). Importantly, this value was found to vary dramatically within Central Hudson's distribution network and was much lower for other distribution nodes. This suggests that the capacity benefits of gas DR are likely to be concentrated in distribution nodes that are experiencing growth or are otherwise operating near capacity limits.

Related to but somewhat distinct from the capacity benefits, gas DR has the potential to increase the reliability and resilience of gas distribution service. Disruptions leading to loss of gas service are relatively rare; however, when they occur, they create serious threats to health and safety and are time consuming and costly to remediate. "The restoration of gas service involves an initial visit to each individual customer to shut off gas valves; work to repair any equipment damage, purge the gas lines, and test for integrity; and a second visit to each individual customer to relight each appliance or manufacturing process and piece of machinery“ (Freeman, 2018). When natural gas service disruptions impact supply to gas power plants, they can also have negative impacts on electricity generation and electricity service. The frequency of these incidents is uncertain and data is generally not publicly available (Freeman, 2018).

Disruptions in gas service can occur for a variety of reasons, both intentional and unplanned, including maintenance or system upgrades, supply constraints, equipment malfunction or error, and damage to facilities from natural disasters or other emergency condition (Argonne, 2002). Under certain types of service disruptions, the flow of gas to a constrained area might be reduced below a normal operating level, but not fully shut off. In this situation, service may be able to be sustained and pressure loss might be avoided if consumption peaks could be smoothed over a longer time period by a demand response program.

Natural Gas Demand Response Programs to Date

While DR programs have become common in the electricity sector, natural gas DR remains relatively underutilized. Only a few utilities have conducted gas DR programs to date. Beginning in 2016/17, SoCalGas implemented a smart thermostat driven and behavior-based Gas DR program. In 2017/18, National Grid began a DR program with 16 commercial and industrial (C&I) customers with large heaters or other natural gas-powered machinery. Con Edison began a thermostat-based gas DR pilot with up to 1,000 residential and commercial customers beginning in winter 2018/19.

Thermostat driven DR programs typically deliver demand reductions over an event period of 2-5 hours. Energy use normally increases in the hours before and after the event period due to pre-conditioning and post-event rebound. Due in part to limitations of typical gas metering infrastructure, public evaluation reports on these programs usually report impacts at the daily level. A recent evaluation of the SoCalGas program is an exception (Bell, et.al., 2019). It includes hourly impact estimates which were produced using interval data from advanced gas meters.

Midwest Utility Gas Demand Response Pilot

During winter 2020/21, a large electric and natural gas distribution utility in the Midwest conducted a gas DR pilot program involving thermostat treatments for residential and small commercial customers. The program evaluation objectives included measuring consumption impacts, customer experience, and identifying factors that influence customer participation; this paper focuses on the findings of the consumption impact analysis. Consumption impact objectives included measurement of pre-heating impacts in the hours preceding events, reductions during event hours, and consumption snapback in the hours following events. These impacts were measured for different customer types and thermostat brands with unique event control algorithms.

The program included an enrolled population of approximately 3,500 residential customers and 400 small or medium business customers (SMB). Residential impacts were measured using a randomized control trial. The SMB customer segment impacts were measured using individual customer baseline models. For both residential and SMB customers, gas consumption was metered at the day interval. The limited temporal resolution raised challenges for measuring the impact of 4-hour DR events. To address this challenge, hourly consumption impacts were measured in heating system runtime minutes using thermostat telemetry data. We then translated DR estimates to hourly gas consumption impacts by estimating the effect of a heating system runtime minute on gas consumption in a second stage model.

The study provides several new insights. It demonstrates the potential for gas demand response in a cold climate Midwest region, the efficacy of experimental and non-experimental methods for gas DR impact measurement, and it provides a solution for measuring gas DR program impacts with limited advanced gas metering infrastructure.

Impact Evaluation Methodology

For both the residential and SMB program segments, we used a two-stage modeling process to estimate the impact of the thermostat demand response treatment during program events. In the first stage, we used thermostat telemetry data to estimate the program impact of the event treatments on customer heating minutes by hour. In the second stage, we estimated the conversion rate of heating minutes to natural gas consumption for customers by sector and thermostat brand. The thermostat telemetry data was provided by the original equipment manufacturers (OEMs) via the program implementer. The telemetry data included hourly resolution temperature setpoints and heating system runtime. Daily resolution gas meter data was provided by the utility.

Despite the consistency in this general approach, there were several important differences in the program implementation and evaluation approaches for the residential and SMB customer segments.

Residential Hourly Heating Minute Impact Methodology

The residential segment was conducted as a randomized control trial (RCT). Within each thermostat brand group, approximately 500 customers were assigned to a control group and the remaining customers were assigned to receive the event treatments. The treatment and control group assignments were re-randomized for each event. We estimated program impacts using a difference-in-differences regression model that compared the treatment and control group heating minutes during event periods.

The regression model included independent variables to account for hourly impacts of temperature (heating degree hours) by customer group, differences in baseline consumption between treatment and control groups on non-event days, and hour-by-date fixed effects. We clustered standard errors by customer account to control for within-customer correlation in natural gas consumption.

We used the following model specification to determine event-specific demand reduction for residential participants:

$$HM_{it} = \sum_{t=0}^{T} \beta_t Datetime_t + \sum_{k=0}^{23} \mu_k Hour_{kt} * I(Treat=1)_i + \sum_{m=1}^{M} \sum_{k=0}^{23} \theta_{mk} Hour_{mkt} * I(Treat=1)_i * I(Event=1)_{mkt} + \sum_{j=1}^{2} \sum_{k=0}^{23} \gamma_{jk} HDH_{it} * I(Treat=1)_i + \varepsilon_{it}$$

Where:

HM_{it}	=	Thermostat heating minutes of customer *i* during date-hour *t*
β_t	=	Datetime fixed effect or average effect of date-hour *t* on customer heating minutes.
$Datetime_t$	=	Indicator variable for date-hour (= 1 if date-hour *t* is the *t*th hour of the sample, *t*=0, 1, 2, …, *T*; = 0 otherwise)
μ_k	=	Average effect of treatment group membership in hour *k* on non-event days
$Hour_{kt}$	=	Indicator variable for hour of the day (= 1 if date-hour *t* is the *k*th hour of the day, k=0, 1, 2, …, 23; = 0 otherwise)
$I(Treat=1)_i$	=	Indicator variable for assignment to treatment group (= 0 otherwise)

θ_{mk}	=	Average impact of the event on treatment customer heating minutes during hour *k* of demand response event *m*
$I(Event=1)_{mkt}$	=	Indicator variable for program event hour (= 1 if date-hour *t* is the *k*th hour, k=1,2,…,J, of event *m*, m=1, 2, …, M)
γ_{jk}	=	Average effect of a heating degree hour on heating minutes for a customer in group *j* (1=treatment, 2=control) in hour *k*
HDH_{it}	=	Heating degree hour for customer *i* in date-hour *t* for a base temperature of 65°F
ε_{it}	=	The error term for customer *i* in date-hour *t*

Small and Medium Business Hourly Heating Minute Impact Methodology

In the SMB segment, the enrolled population was smaller (< 500 total) and we anticipated greater heterogeneity in customer heating patterns due to differences in business schedules. For these reasons, the SMB segment was not conducted as an RCT. Instead, all enrolled participants received the event treatments. To measure program impacts, we used regression methods with non-event day data to predict counterfactual baseline heating.

We constructed a series of candidate predictive models using non-event day thermostat data. We evaluated the predictive accuracy of each model on a test set of non-event days when temperature conditions were similar to event days. The candidate models included a variety of temperature condition variables (such as temperature, heating degree hours, or heating degree hour buildup) and various datetime variables (such as day of week or hour of day). For each facility, we selected the candidate model that produced the most accurate prediction (minimum root mean squared error) for the final impact estimation.

We used the selected models to predict counterfactual baseline heating minutes. Finally, we compared the counterfactual to actual heating minutes on event days to estimate program impacts in each hour. Table 1 lists the candidate models and the number of facilities for which each model was selected.

Table 1. SMB Candidate Models

Candidate Model	Independent Variables	Number of Facilities
1	Hour	32
2	Hour, Day of Week	10
3	Hour, Week	13
4	Hour, Day of Week, Week	14
5	Hour, HDH25[a]	8
6	Hour, HDH50[b]	9
7	Hour, HDH25 Buildup[c]	8
8	Hour, HDH50 Buildup	8
9	Hour, HDH25, HDH50 Buildup	9
10	Hour, HDH50, HDH25 Buildup	11
11	Hour, Week, HDH25	5
12	Hour, Week, HDH50	14

Candidate Model	Independent Variables	Number of Facilities
13	Hour, Week, HDH25 Buildup	9
14	Hour, Week, HDH50 Buildup	6
15	Hour, Week, HDH25, HDH50 Buildup	9
16	Hour, Week, HDH50, HDH25 Buildup	12
17	Weekday, HDH25	5
18	Weekday, HDH50	8
19	Weekday, HDH25 Buildup	4
20	Weekday, HDH50 Buildup	2
21	Weekday, HDH25, HDH50 Buildup	6
22	Weekday, HDH50, HDH25 Buildup	4
23	Hour*Weekday, HDH25	6
24	Hour*Weekday, HDH50	10
25	Hour*Weekday, HDH25 Buildup	6
26	Hour*Weekday, HDH50 Buildup	7
27	Hour*Weekday, HDH25, HDH50 Buildup	8
28	Hour*Weekday, HDH50, HDH25 Buildup	16
29	Hour, Weekday, HDH25	9
30	Hour, Weekday, HDH50	8
31	Hour, Weekday, HDH25 Buildup	9
32	Hour, Weekday, HDH50 Buildup	5
33	Hour, Weekday, HDH25, HDH50 Buildup	3
34	Hour, Weekday, HDH50, HDH25 Buildup	6
35	Hour, Hour*HDH25	29
36	Hour, Hour*HDH50	44
37	Hour, Hour*HDH25 Buildup	18
38	Hour, Hour*HDH50 Buildup	36
Total		**426**

[a] Heating degree hour for a base temperature of 25°F
[b] Heating degree hour for a base temperature of 50°F
[c] Cumulative function of heating degree hours over preceding 24 hours

Heating Minutes to Gas Consumption Conversion Rate Estimation

In a second stage model, we estimated the conversion rate between heating minutes and natural gas consumption. We merged metered natural gas consumption and thermostat telemetry data to produce a table with average hourly heating minutes for each metered natural gas consumption interval. Then, we used a regression model to estimate the average effect of a heating minute on natural gas consumption. For residential participants, we estimated a single conversion rate for each brand. The analysis was segmented by brand to control for any potential correlation in thermostat brand choice and heating equipment choice. However, the results were very similar across brands: one minute of heating caused an increase in gas consumption of approximately 0.95 cubic feet.

For the SMB segment, we stratified participants into groups based on median daily natural gas consumption to account for differences in heating system capacities (i.e., MBH). We estimated a separate conversion rate for each group; these rates are displayed in Table 2. There was a strong positive relationship between median hourly natural gas consumption and natural gas consumption per heating minute.

Table 2. SMB Conversion Rates by Consumption Bin

Median Daily Gas Consumption (cf)	cf per Heating Minute
Less than 25	0.91
25 to 49	1.14
50 to 99	1.68
100 to 199	3.23
200 to 299	4.60
300 and greater	6.08

Due to limitations of the thermostat telemetry data for brand C, we were unable to estimate a conversion rate for brand C participants directly. For residential brand C participants, we applied the average of the conversion rates estimated for participants with brands A and B, 0.97 and 0.93 cf per heating minute, respectively. For SMB brand C participants, we applied the conversion rates from the corresponding median consumption strata, of 1.52 cf per heating minute.

Impact Evaluation Results

During the 2020-2021 winter event season, the utility called 10 natural gas demand response events. The events occurred on non-holiday weekdays between January 1, 2021 and February 28, 2021, under cold temperature conditions (less than 23°F), corresponding with periods of high natural gas demand. Table 3 summarizes the program performance across all 2021 events.

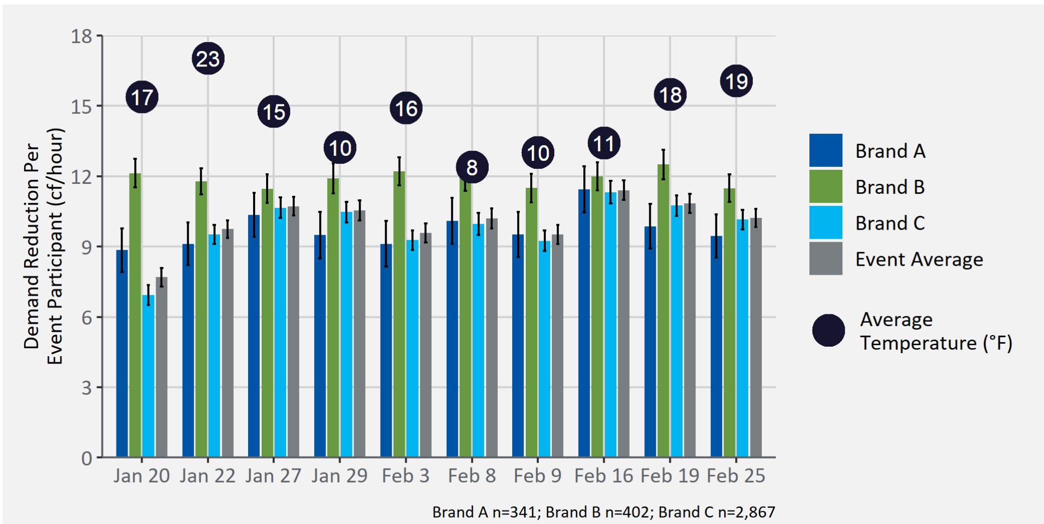
Table 3. Event Summary

Date	Event Time	Event Participants[a]	Average Temperature[b]	Program Average Demand Reduction (Mcf/hour)[c]
January 20	6 a.m. - 10 a.m.	3,853	17°F	30.8
January 22	6 a.m. - 10 a.m.	3,831	23°F	38.0
January 27	6 a.m. - 10 a.m.	3,943	16°F	42.4
January 29	6 a.m. - 10 a.m.	3,965	10°F	42.7
February 3	6 a.m. - 10 a.m.	3,976	16°F	38.2
February 8	6 a.m. - 10 a.m.	3,982	8°F	42.3
February 9	6 a.m. - 10 a.m.	3,943	10°F	40.4
February 16	5 a.m. - 9 a.m.	4,067	11°F	46.6
February 19	5 a.m. - 9 a.m.	4,046	18°F	44.7
February 25	5 a.m. - 9 a.m.	4,005	20°F	42.5
Average	**-**	**3,961**	**15°F**	**40.9**

[a] Residential and SMB customers whose thermostats were controlled during the event.
[b] Weighted average of local temperature conditions recorded by each participant's smart thermostat.
[c] Aggregate program demand reduction (residential and SMB) calculated as the average hourly demand reduction across the four hours of the event.

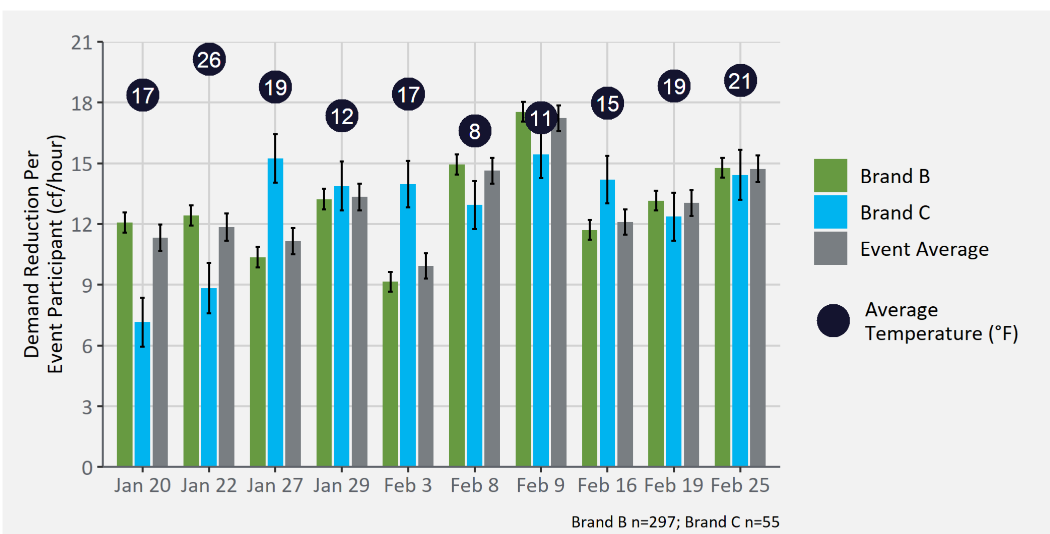
Peak Demand Reduction

The first seven events conducted in January and February ran from 6 a.m. to 10 a.m. The final three program events conducted in February began at 5 a.m. and ran until 9 a.m. All program events lasted four hours. Figure 1 and Figure 2 depict the per event participant demand reduction for each thermostat brand for residential and SMB customers, respectively.



Note: Error bars depict 90% confidence bounds.

Figure 1. Average Per-Participant Demand Reduction by Thermostat Brand – Residential Sector



Note: Error bars depict 90% confidence bounds.

Figure 2. Average Per-Participant Demand Reduction by Thermostat Brand – SMB Sector

The program achieved statistically significant demand reduction among all thermostat brands in both sectors during each program event. Among residential participants, the average per-customer hourly load reduction ranged from 7.7 cf on January 20 to 11.4 cf on February 16. The maximum hourly per-participant demand reduction of 20.8 cf occurred on February 16 between 5 a.m and 6 a.m. SMB customers' average hourly per-participant impacts ranged between 9.9 cf on February 3 and 17.2 cf on February 9. The maximum hourly impact of 26.8 cf was achieved between 6 a.m and 7 a.m. during the February 9 event. Program-level average hourly impacts ranged from 30.8 Mcf to 46.6 Mcf.

The average reduction in heating system runtime during event hours was approximately 50% for residential participants and 34% for SMB participants. By brand, the residential demand reductions were 39.0%, 52.2%, and 51.8% for brands A, B, and C, respectively. The SMB demand reductions were 33.1% and 41.0% for brands B and C, respectively.

Pre-Conditioning and Post-Event Rebound Demand Impacts

In addition to demand reduction during the four hours of each program event, the thermostat treatment algorithm affected natural gas consumption in the hours preceding and following each event. Figure 3 depicts the average hourly heating demand for residential treatment and control customers on February 19, 2021. An increase in treatment group heating consumption due to thermostat pre-conditioning is visible beginning three hours preceding the event. Following the event, there is a visible rebound effect as the treatment group customers consume more natural gas than the control group customers. The increase in consumption following peak events can be attributed to increased heating as the participants' homes return to a non-event thermostat setpoint. This effect is commonly known as rebound. Other demand response programs see similar effects following events. Rebound impacts persisted for three hours following the event, but the majority of the rebound effect occurred in the first post-event hour.

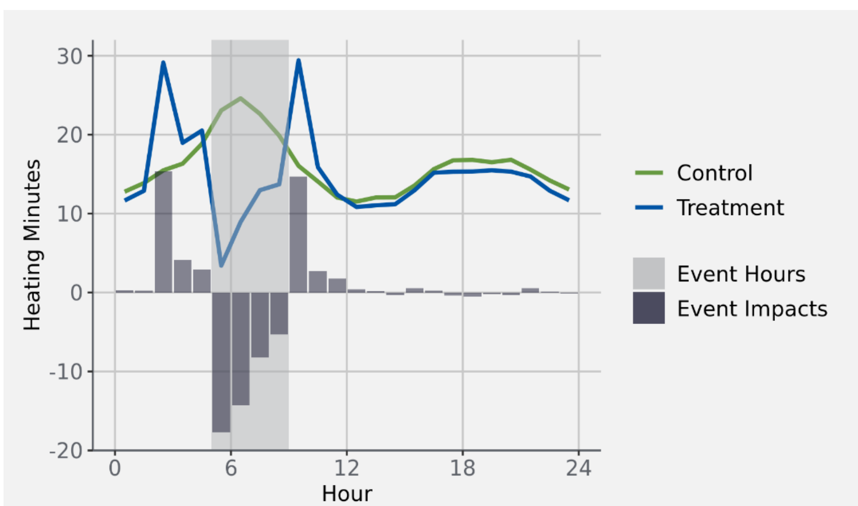


Figure 3. Residential Participant Heating Minute Impacts – February 19, 2021

There were significant increases in demand before and after each event due to the pre-conditioning component of the event treatment and to post-event rebound. Although pre-conditioning and rebound impacts varied slightly by thermostat brand, the largest demand impacts occurred during the three hours immediately preceding and following each event.

Pre-conditioning demand impacts for the three hours preceding an event averaged 6.1 cf per event per participant (24.0 Mcf aggregate per event impact) across all events. Average post-event rebound impacts were 5.5 cf per event per participant (21.7 Mcf aggregate per event impact).

After accounting for event hour impacts, pre-conditioning impacts, and post-event rebound impacts, customers saved an average of 4.8 cf on event days. Increased pre- and post-event heating energy use offset 71% of the event's energy use reduction. The program produced average aggregate energy savings of 19.0 Mcf per event.

Conclusions

Pilot programs can help determine the technical feasibility and performance potential for natural gas demand response. This evaluation found that, over a 4-hour event period, thermostat gas DR can reduce customer heating system runtime by 50% and 34% for residential and SMB customers, respectively. We also found that the limitations of evaluation with daily gas meter data can be overcome by measuring hourly program impacts with thermostat telemetry data. Measuring hour interval impacts is particularly important for thermostat programs with impacts that are focused in a few target hours. Additional research is needed to determine whether thermostat-driven gas demand response can address capacity constraints and provide resilience and reliability benefits cost-effectively.

References

Argonne National Laboratory Infrastructure Assurance Center (2002). "Planning for Natural Gas Disruptions." https://publications.anl.gov/anlpubs/2003/02/45798.pdf

Bell, Eric, Stephanie Bieler, Ty Lehman, Nicholas Cain (2019). "2018-2019 Winter Load Impact Evaluation of SoCalGas Smart Therm Program." Nexant.http://www.calmac.org/publications/SoCalGas_2019_DR_Evaluation_Report_-_PUBLIC_FINAL.pdf

Bode, Josh, Adriana Ciccone, and Ross Beppler (2020). "Central Hudson Location-Specific Avoided Gas Distribution Costs Using Probabilistic Forecasting and Planning Methods." Demand Side Analytics. https://documents.dps.ny.gov/search/Home/ViewDoc/Find?id=%7B930B600C-29C2-453E-A505-C6F14C2BDB75%7D&ext=pdf

Freeman, Gerad, Jay Apt, and Michael Dworkin (2018). "The Natural Gas Grid Needs Better Monitoring." *Issues in Science and Technology* 34, no. 4. https://issues.org/the-natural-gas-grid-needs-better-monitoring/

Sato, IIchiro, Beth Elliott, and Clea Schumer (2021). "What is Carbon Lock-in and How Can We Avoid It?" World Resources Institute. https://www.wri.org/insights/carbon-lock-in-definition