

Roth, Alexander; Schmidt, Felix

Article — Accepted Manuscript (Postprint)

Not only a mild winter: German consumers change their behavior to save natural gas

Joule

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Roth, Alexander; Schmidt, Felix (2023) : Not only a mild winter: German consumers change their behavior to save natural gas, Joule, ISSN 2542-4351, Elsevier, Amsterdam, Vol. 7, Iss. 6, pp. 1081-1086,
<https://doi.org/10.1016/j.joule.2023.05.001>

This Version is available at:

<https://hdl.handle.net/10419/296667>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

This is the postprint of the article Not Only a Mild Winter: German Consumers Change Their Behavior to Save Natural Gas : Commentary. Published in Joule 7 (2023), 6, S. 1081–1086, available online at: <https://doi.org/10.1016/j.joule.2023.05.001>

Not only a mild winter: German consumers change their behaviour to save natural gas

Alexander Roth^{a,*}, Felix Schmidt^a

^a*German Institute for Economic Research (DIW Berlin), Mohrenstraße 58, 10117 Berlin, Germany*

Keywords: natural gas consumption, energy savings, behavioural changes

1. Introduction

By the start of the 2022/2023 heating season, Germany and many other European countries found themselves facing a potential gas supply shortage in the wake of Russia's invasion of Ukraine. In search of a response, authorities called on residential and commercial sectors to save natural gas. Exploiting winter 2022/23 as a “natural experiment”, we shed light on the magnitude of behavioural gas savings using open data and a machine learning method. Despite being exposed to incomplete price signals, we find significant behavioural gas savings by German households and businesses, contributing to closing the supply gap. We uncover temperature-dependent saving dynamics and discuss the potential roles of different drivers of this change. Finally, we highlight the pivotal role of a timely and continuous provision of openly accessible data and analysis to inform the general public as well as policymakers.

*Corresponding author

Email address: aroth@diw.de (Alexander Roth)

2. Context

The Russian invasion of Ukraine in February 2022 has created an unprecedented supply crunch in European natural gas markets. Up until February 2022, Russia had been Europe's largest supplier of natural gas, expanding its position in prior years. Doubting the reliability of Russia's gas supplies, the question of whether enough gas would have been supplied to the European market led to spiralling wholesale gas prices. At the end of August 2022, prices peaked at over 300 Euro per megawatt hour (MWh) at the benchmark hub TTF after Russia stopped delivering gas through its Nord Stream 1 pipeline.^[1] Slowly rising in the months prior to the invasion, prices had been fluctuating around 20 Euro per MWh in recent years.^[1] Following the closure of Nord Stream 1, the security of supply was called into question with respect to the upcoming winter of 2022/23.^[2]

Within a year, (Central) Europe's gas supply structure changed radically. While historically, around 40% of all gas imported to Germany had been coming through Russian pipelines, this number dropped to almost 0% by the end of 2022.^[3] Much of the Russian supply was substituted by additional pipeline imports from Norway and liquefied natural gas (LNG) shipments from other countries. The remaining potential shortfall gave rise to a discussion on how much gas could and would be saved by whom.

With respect to gas consumption, there are three principal groups: gas-fired power plants, large industrial consumers, and the residential and commercial sectors, which comprise households and small- and medium-sized businesses. Gas-fired power plants consume gas for electricity production, yet some also supply heat to district

1
2
3
4
5
6 heating networks. Large industrial consumers use gas either as feedstock or source
7 of process heat. The residential and commercial sectors need gas predominantly to
8 satisfy heat demand.
9

10
11
12 These consumer groups are different in terms of the price signals they receive,
13 as well as the potential for and consequences of gas demand reductions or enforced
14 curtailment. Gas-fired power plants usually buy gas short-term to serve peak elec-
15 tricity demand and thus react immediately to price signals in both electricity and
16 gas markets. Provided there is sufficient alternative electricity supply, e.g. from
17 coal-fired power plants, gas demand from the power sector is rather flexible. Large
18 industrial consumers, unless protected by long-term gas supply contracts or com-
19 prehensive hedging, are similarly exposed to price changes in the spot market and
20 therefore have an incentive to reduce gas consumption in case of a supply crunch.
21 At least in the short run, the industry can reduce its gas consumption by curbing
22 production, substituting the energy carrier, or buying alternative upstream products.
23 Mostly supplied under fixed-price contracts, residential and commercial consumers
24 do not bear the consequences of rising prices in the spot market until a contract has
25 to be renewed. Even in the case of an acute gas shortage, it is not clear whether
26 a controlled gas curtailment of supply to residential and commercial sectors in the
27 distribution grids would have been possible, as it would have been challenging to
28 implement for various technical^[4,5] and political reasons.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 In the face of a looming gas shortage, the public debate initially concentrated
49 on industry halting production, leading to a strong economic downturn, the size of
50 which was debated controversially among economists.^[6,7] To avoid dire economic
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

consequences of production cutbacks of industrial consumers and because of limited means for the government to impose rationing, voluntary savings by residential and commercial sectors eventually gained importance in closing the gas supply gap.

3. Gas savings from changes in behaviour

Since the beginning of the gas supply crunch, Germany has been the focus of discussion due to its large economy and relatively high dependence on Russian gas imports. In September 2022, the German Federal Network Agency *Bundesnetzagentur* announced that a 20% reduction in gas consumption (compared to the average consumption of the preceding four years) would have been necessary to avoid an acute gas shortage.^[8]

In the following, we aim to shed light on the efforts by residential and commercial sectors to save gas. The strong dependency of residential and commercial gas demand on weather conditions implies that relatively warmer or colder weather has a large effect on whether the target is actually achievable or not. Building on a rich literature on the relationship between heat demand, gas demand, temperatures and prices^[9–12], we use a very flexible machine learning method to isolate those gas demand drivers that are not governed by weather variations. We subsume these drivers as the *behavioural component*.

The method used in this commentary to estimate savings is a causal forest, which has two important features: (1) It is fully non-parametric and data-driven, and (2) it allows isolating savings effects differentiated by temperature. Causal forests^[13]

extend a classic machine learning algorithm, random forests^[14]. The general idea of random forests is to partition the data set based on values of explanatory variables and fit local models within these partitions, which are together capable of representing non-linear relationships without having to specify a functional form. Causal forests extend this concept by using the same logic as a tool to identify local saving effects. We provide extensive explanations, details, and robustness checks of our model in the Supplemental Information sections SI.2-SI.4. The causal forest model enables us to *predict* daily behavioural savings depending on the weather conditions of the day. In order to control for weather conditions, we include mean, minimum and maximum temperatures of a given day as well as several lags to control for thermal inertia. Irradiation effects are proxied by sunshine duration, and we include month and weekend/holiday indicators to account for behavioural variations.

Our model allows us to recover two alternative scenarios of *estimated* consumption. The first scenario is the estimated *actual* consumption, including behavioural savings. The second scenario is the estimated *counterfactual* consumption, which would be expected in the absence of the savings. By design, the difference between these two scenarios yields our estimate of *behavioural savings*. By focusing on estimated counterfactual consumption and estimated actual consumption (instead of *observed* consumption), we ensure a like-for-like comparison and that our savings are not driven by random error. This assumes implicitly that the model errors, given by the difference between the estimated actual consumption and the observed consumption, would have been the same in the absence of behavioural savings.

In the upper panel of Figure 1, the estimated actual consumption is depicted

as a solid black line, while a solid red line represents the estimated counterfactual consumption (in the absence of savings). The dashed black line gives the observed consumption. We start measuring the savings effect as of September 2022, when the risk of a supply shortage became pressing with the start of the heating period and the end of Nord Stream deliveries. Nonetheless, our model allows for the possibility of behavioural savings from the beginning of the Russian invasion of Ukraine on 24 February 2022. We discuss the implications of this assumption in detail in the Supplemental Information section SI.4.

Gas consumption has been going up as expected with colder temperatures (Figure 1). With the beginning of the heating season in September, we see that German residential and commercial sectors have consistently saved between 66 and 285 GWh of gas per day. As revealed in the lower panel of Figure 1, estimated savings are statistically significant for all days in the September to December period. December 2022 was exceptionally cold, also reflected by spiking gas demands. Around the Christmas period, savings efforts diminished. Cumulatively, we estimate that households and commercial sectors have saved ca. 23 TWh [95% CI: 18.7; 27.3] by changing their behaviour from the beginning of September until the end of December 2022.

Relying on the results above, we can attribute the differences in gas consumption between 2022 and the average of the period 2018-2021 to different effects (Figure 2). The weather effect (grey) is computed as the difference between the estimated 2018-2021 average consumption and the estimated counterfactual consumption in 2022. Behavioural savings (red) result from the difference between estimated actual and counterfactual consumption. The sum of weather and behavioural savings does

not add up to the total difference in consumption, represented by the solid line, due to the unobserved error component discussed above. The 20% savings target defined by German Federal Network Agency is reflected by the dashed line.

Total savings compared to the average of 2018-2021 varied substantially between different weeks (Figure 2). This variation is mostly driven by the weather component. Meanwhile, the behavioural component remains relatively stable, slightly increasing over time. Compared to 2018-2021, we observe two cold spells: one in September (as of calendar week 36) and one in mid-December (as of calendar week 50), in which the weather component drove up gas consumption. Even in these colder periods, estimated behavioural savings did not change much. In the last two weeks of the year, savings decreased slightly compared to the previous weeks. This may be explained either by the Christmas period or by a reduced urgency, as it became increasingly evident by December that a gas shortage in the winter of 2022/23 would be rather unlikely. Gas storage levels remained well above the range of previous years.

On aggregate, we find that the weather effect alone did not play a significant role when comparing the September to December 2022 gas consumption with previous years (right panel of Figure 2). At least for the first half of the winter, this is possibly at odds with other analyses asserting that a comparably mild winter induced most savings.^[15] Consistent behavioural savings contrast highly variable weather-related savings. Especially the cold spell in December offset most of the savings by weather due to milder temperatures in the weeks before. However, the weather may have had an indirect effect, as a colder winter would have made it even harder for households to save gas in the same way.

The winter months of 2022 also shed light on the savings dynamics of the residential and commercial sectors relative to temperatures. We find a negative relationship between relative gas savings, defined as absolute gas savings divided by estimated counterfactual consumption, and temperature (lower panel, Figure 3). The residential and commercial sectors seem to relatively easily suppress their heating demand when temperatures are rather mild. These levels of relative savings cannot be carried over to lower temperatures. If outside temperatures are around 12°C, decreasing heating efforts by a certain amount will have a much lower effect on room temperatures compared to a situation when outside temperatures range around 0°C.

Regarding the relevance of averting a gas shortage, relative savings are, however, only of minor importance. Therefore, we highlight the substantial and consistent absolute savings during cold temperature days (upper panel of Figure 3). Although they fell short of the targeted 20% goal by the federal regulator, they added more to averting a gas shortage than the higher relative savings in autumn.

4. Conclusions and outlook

Winter 2022/23 happened to be a “natural experiment” for Europe and Germany on how the economy would react to a gas supply crunch or even a looming shortage. It tested the capacity and willingness of households and commercial consumers to cut gas demand mainly used for heating. Using a data-driven causal forest model, we can show that residential and commercial sectors have reduced their gas consumption. In contrast, the weather had even an increasing effect.

1
2
3
4
5
6
7 The reasons for these savings could be manifold, including but not limited to
8 increased prices, clear communication by officials, changed expectations, and political
9 conviction and solidarity.

10
11
12 As most of Germany's residential and commercial sectors face fixed price regimes,
13 wholesale market price spikes usually do not affect consumers directly. Short-lived
14 price hikes on the wholesale market typically do not translate into higher long-term
15 retail tariffs. For the prolonged price increase in the wake of the Russian invasion
16 of Ukraine, average retail prices only reacted sluggishly^[11]. Furthermore, staggered
17 contractual periods and the unavailability of individual-level consumption data make
18 it challenging to compute precise price elasticities at the retail level. Notwithstand-
19 ing, higher prices have certainly affected the estimated behavioural savings. Yet, the
20 precise impact of prices on German residential and commercial sectors remains, for
21 the moment, opaque.

22
23
24 As we observe savings despite incomplete price signals, we suggest they might
25 have also been driven by a response to public communication. As September came
26 to an end, Germany had experienced a colder start into autumn than usual, and the
27 German Federal Network Agency, *Bundesnetzagentur*, and its president urged res-
28 idential and commercial sectors to reduce consumption. Consequently, the agency
29 released the aforementioned target of a 20% demand reduction. The president re-
30 peated this plea several times. In addition, consumers could have saved addition-
31 ally in expectation of higher prices. Clear communication by the Federal Network
32 Agency raised public awareness of the role of storage levels and their effect on whole-
33 sale prices and, eventually, contract prices. Consumers are likely to have understood

1
2
3
4
5
6 that lower consumption levels today would keep storage levels sufficiently high in
7
8 order to avoid costly additional imports. Other reasons might have played a role as
9
10 well. Some consumers could have regarded saving gas as a part of responsible civil
11
12 behaviour. Political beliefs towards the support of Ukraine (or Russia) could also
13
14 have (de-)motivated the savings behaviour of some households.
15
16

17 Importantly, we want to highlight the essential role of continuous and timely
18
19 data provision and analysis for public debate and policymaking. Transparency and
20
21 publicly available data are crucial for consumers and policymakers, not only to bet-
22
23 ter understand the topic but also to track whether measures and their efforts have
24
25 any effect. In autumn 2022, little publicly available evidence existed on whether
26
27 and how strong the residential and commercial sectors would help in savings gas to
28
29 avoid a potential gas shortage in the winter months. Several platforms began to
30
31 publish analyses on various aspects of the energy crunch, such as consumption data,
32
33 storage levels, prices etc. On the “Open Energy Tracker”^[3], we have been tracking
34
35 behavioural gas savings of residential and commercial sectors since October 2022,
36
37 providing the public with timely insights. The results and methods in this commen-
38
39 tary are based on those published in a less elaborate form on the “Open Energy
40
41 Tracker”.
42
43

44 Despite the impact that data and analyses might have already had on policy
45
46 and consumer behaviour in this gas crisis, improved data quality, e.g. by means of
47
48 an accelerated smart meter roll-out, could yield further benefits. It could enhance
49
50 the quality of the analysis by uncovering drivers of consumer behaviour and thereby
51
52 increase the policy relevance of real-time analyses. It could also allow for more
53
54
55
56

1
2
3
4
5
6 direct pricing mechanisms that prompt an immediate consumer response to wholesale
7 market developments.
8
9

10 Finally, all results in this piece can only be regarded as a snapshot in time, and a
11 complete picture will only emerge in a continued analysis. The estimates presented
12 in this commentary will be continuously updated online.^[3] We believe that with a
13 data-driven analysis of events, the public and policymakers have an important tool
14 at hand to assess the success of saving efforts and their policies.
15
16
17
18
19
20
21
22

23 **Acknowledgments**

24
25
26 We thank Wolf-Peter Schill for his very helpful comments and remarks. This work
27 benefited from a research grant by the German Federal Ministry of Education and
28 Research (BMBF) via the Kopernikus project Ariadne (FKZ 03SFK5N0).
29
30
31
32
33
34

35 **Author contributions**

36 Conceptualization: AR, FS; Methodology: AR, FS; Software: FS; Formal analysis
37 FS; Investigation: AR, FS; Data Curation: AR, FS; Writing - Original Draft: AR,
38 FS; Visualization: AR, FS
39
40
41
42
43
44
45
46

47 **Declaration of interests**

48 The authors declare no competing interests.
49
50
51
52
53
54
55
56

References

- [1] TradingEconomics.com, EU Natural Gas (2023).
URL www.tradingeconomics.com/commodity/eu-natural-gas
- [2] M. Murphy, Nord Stream 1: Russia shuts major gas pipeline to Europe, BBC News (Sep 2022).
URL www.bbc.com/news/world-europe-62732835
- [3] W.-P. Schill, A. Roth, Open Energy Tracker (2023).
URL www.openenergytracker.org/en/
- [4] R. Winkelhahn, Gasmangelsicherung: Können Heizungen bei Gasmangel im Winter einfach ausfallen?, Handelsblatt (Sep. 2022).
URL www.handelsblatt.com/unternehmen/energie/gasmangelsicherung-koennen-heizungen-bei-gasmangel-im-winter-einfach-ausfallen/28706050.html
- [5] J. Haase, Potsdam spart bislang kaum Energie: Stundenweise Stromabschaltung nicht vorgesehen – aber nicht ausgeschlossen, Der Tagesspiegel Online (May 2022).
URL www.tagesspiegel.de/potsdam/landeshauptstadt/potsdam-spart-bislang-kaum-energie-zeitweise-stromabschaltung-nicht-vorgesehen-aber-nicht-ausgeschlossen-8710945.html
- [6] R. Bachmann, D. Baqaee, C. Bayer, M. Kuhn, A. Löschel, B. Moll, A. Peichl, K. Pittel, M. Schularick, What if? The Economic Effects for Germany of a Stop of Energy Imports from Russia, EconPol Policy Report 36 6 (2022).
URL www.econpol.eu/publications/policy_report_36
- [7] T. Krebs, Economic Consequences of a Sudden Stop of Energy Imports: The Case of Natural Gas in Germany, ZEW - Centre for European Economic Research Discussion Paper No. 22-021 (2022).
URL www.ssrn.com/abstract=4168844

- [8] Bundesnetzagentur, Aktuelle Lage Gasversorgung - Gasverbrauch der Haushalte steigt im Moment zu stark an (Sep. 2022).
URL https://www.bundesnetzagentur.de/SharedDocs/Pressemitteilungen/DE/2022/20220929_Verbrauchsdaten.html;jsessionid=3889BF046FB5D7C13E94CBCB91195CCA?nn=1077982
- [9] A. Henley, J. Peirson, Non-linearities in electricity demand and temperature: Parametric versus non-parametric methods, *Oxford Bulletin of Economics and Statistics* 59 (1) (1997) 149–162. doi:10.1111/1468-0084.00054.
- [10] K. Wojdyga, An influence of weather conditions on heat demand in district heating systems, *Energy and Buildings* 40 (11) (2008) 2009–2014. doi:10.1016/j.enbuild.2008.05.008.
- [11] O. Ruhnau, C. Stiewe, J. Muessel, L. Hirth, Gas demand in times of crisis: energy savings by consumer group in Germany, *ZBW - Leibniz Information Centre for Economics Preprint* (261082) (2022).
URL <https://ideas.repec.org/p/zbw/esprep/261082.html>
- [12] C. Bantle, J. Wiersich, Gasverbrauch: Heizen wir weniger als sonst?, *Tech. rep.*, Bundesverband der Energie- und Wasserwirtschaft (Sep. 2022).
URL www.bdew.de/media/documents/221209_Diskussionspapier_Gaseinsparung_FINAL.mitAP
- [13] S. Wager, S. Athey, Estimation and Inference of Heterogeneous Treatment Effects using Random Forests, *Journal of the American Statistical Association* 113 (523) (2018) 1228–1242. doi:10.1080/01621459.2017.1319839.
- [14] L. Breiman, Random Forests, *Machine Learning* 45 (1) (2001) 5–32. doi:10.1023/A:1010933404324.
- [15] J. Blas, Analysis | The New European Energy Normal Remains Rather Painful, *Washington Post* (Mar. 2023).
URL www.washingtonpost.com/business/energy/the-new-european-energy-normal-remains-rather-painful/2023/03/06/67bf3d84-bbdf-11ed-9350-7c5fccd598ad_story.html

Figure legends

Figure 1

Title: Daily actual and counterfactual gas consumption

Description: The upper panel (**A**) shows the modelled actual and counterfactual daily gas consumption paths from September 2022 - December 2022. The solid line in the lower panel (**B**) gives the estimated daily behavioural savings (corresponding to the shaded area in the upper panel). The dashed lines define the 95% confidence interval of the estimated savings.

Figure 2

Title: Gas savings disaggregated into weather and behavioural components vs 2018-21 average

Description: Gas savings in 2022 compared to 2018-21 consumption, disaggregated into behavioural (red) and weather component (grey). The solid black line represents total savings and the dashed line the savings goal of 20%. The *left panel* (**A**) provides a weekly view, while the *right panel* (**B**) shows the accumulated savings for the calendar weeks 36-52 2022.

Figure 3

Title: Relationship between behavioural savings and mean temperature

Description: This figure shows the relationship between estimated behavioural savings and the mean temperature. The *upper panel* (**A**), shows absolute savings in GWh, while the *lower panel* (**B**) displays relative behavioural savings defined as absolute savings divided by estimated counterfactual consumption.

Figure 1 main body

[Click here to access/download;Figure;Figure 1.pdf](#)

— Exp. actual consumption — Exp. counterfactual consumption - - Observed consumption

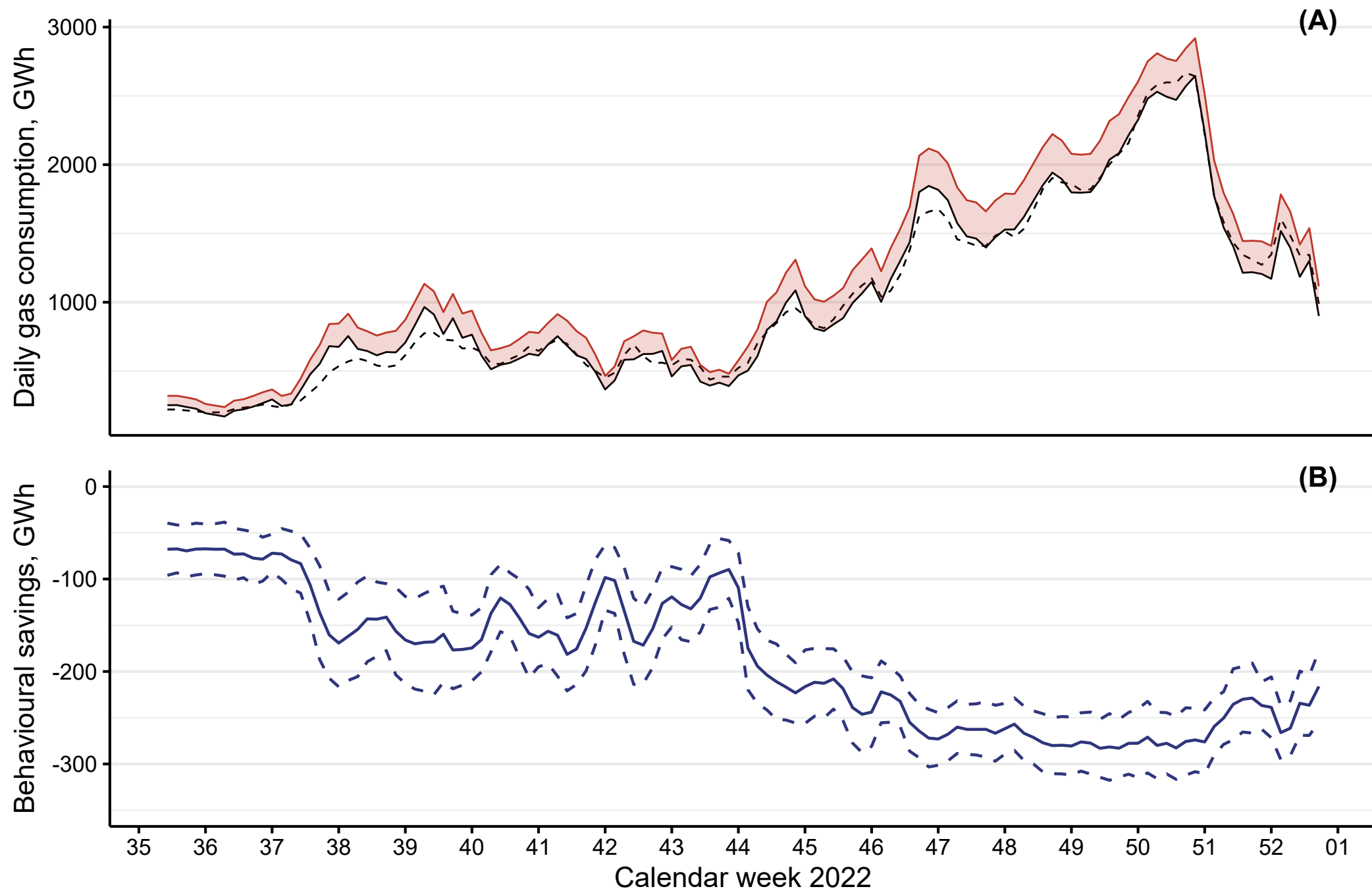
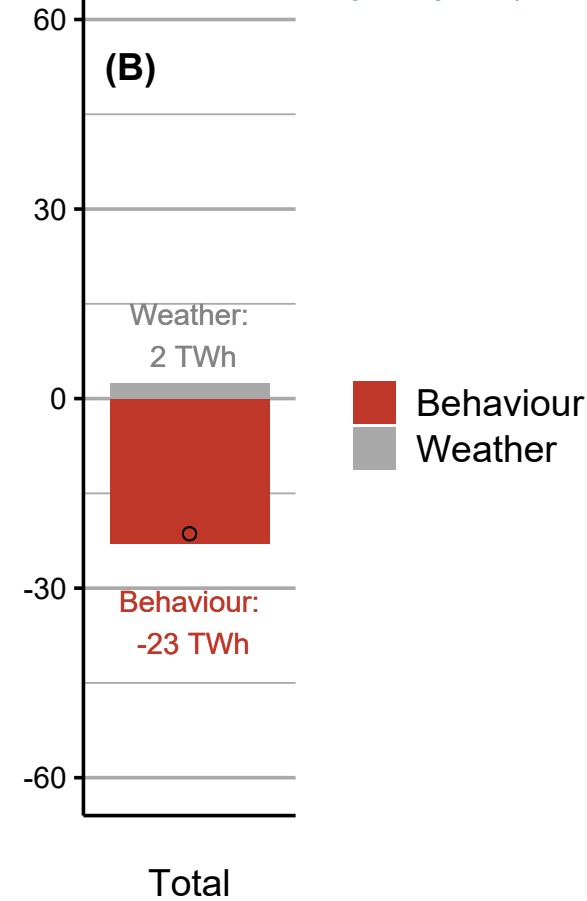
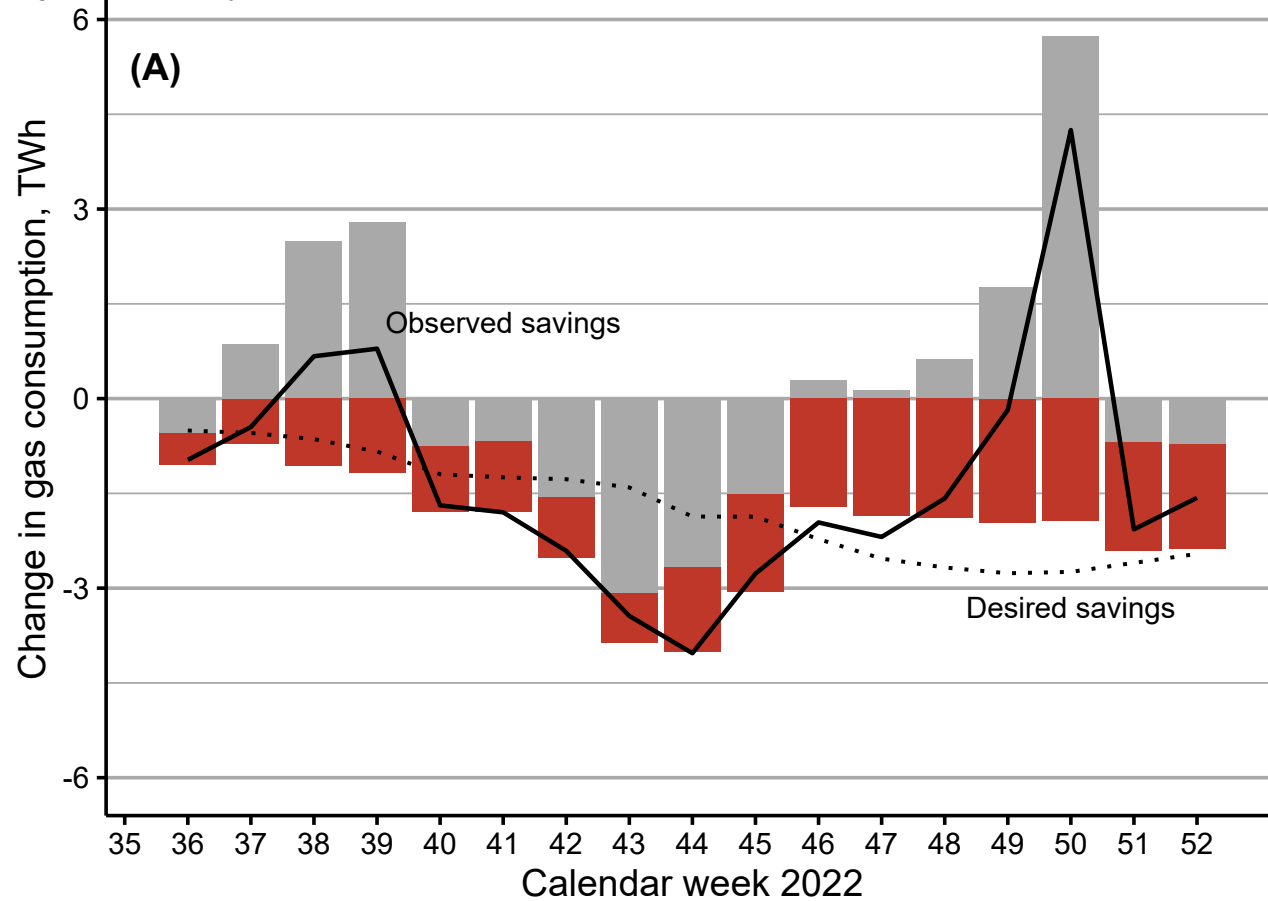
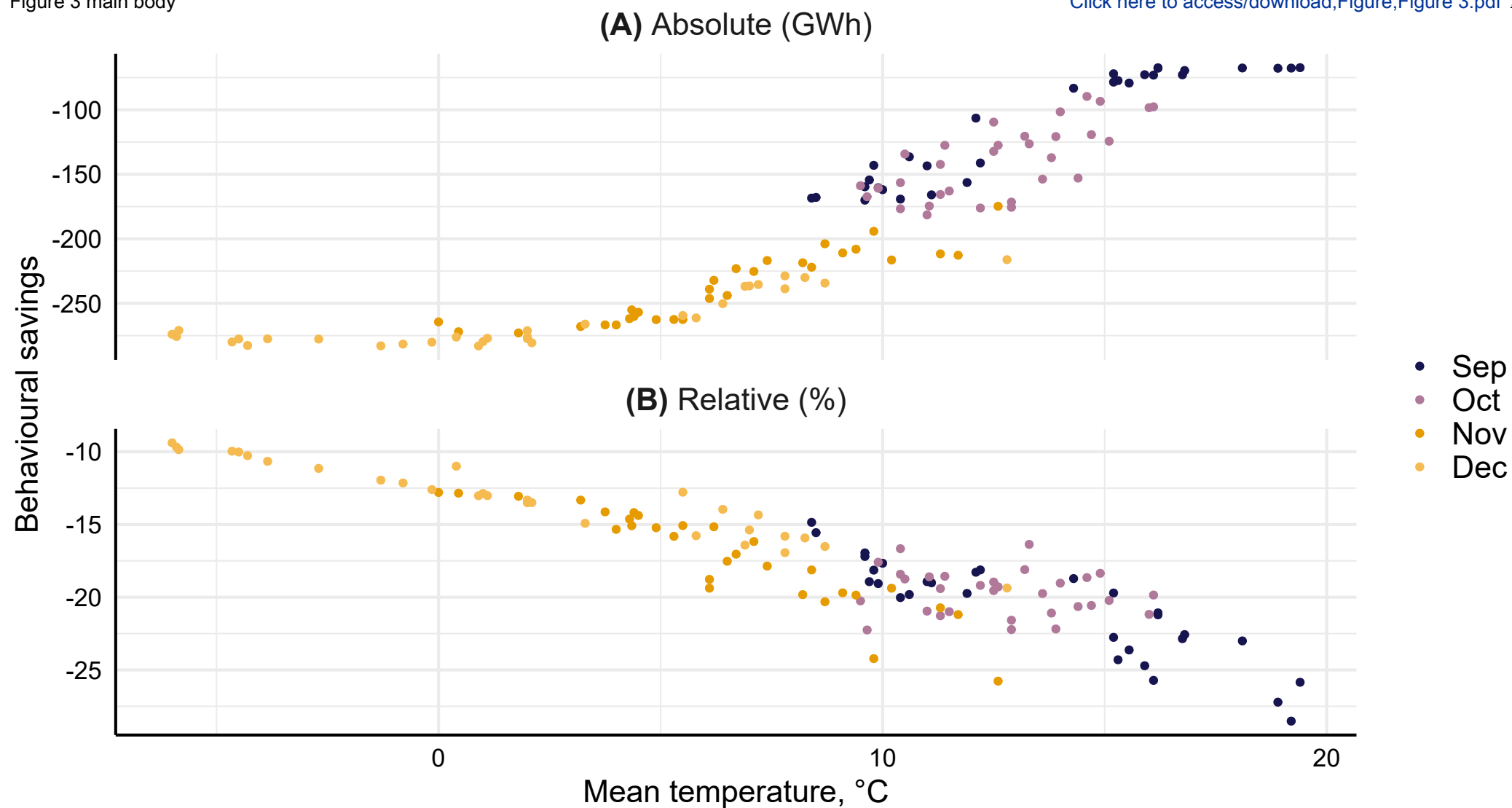


Figure 2 main body

[Click here to access/download;Figure;Figure 2.pdf](#)





SI. Supplemental Information

This section provides details on the deployed methodology and presents supplemental results. Section SI.1 discusses data sources and provides some summary statistics. Section SI.2 gives a methodological introduction to local linear causal trees. Section SI.3 contrasts it with alternative approaches. Section SI.4 provides supplemental results on robustness.

SI.1. Data and descriptive statistics

SI.1.1. Gas consumption data

Instantaneous gas consumption metering for residential and commercial customers is still rare in Germany, such that accurate day-by-day consumption profiles for individual households or business units are unavailable [1]. In the absence of directly metered data, the German Network Agency relies on residual load data published by the German gas exchange Trading Hub Europe (THE). The residual load is derived by taking the difference between gas inflows and gas outflows from the network to downstream networks, storages, other countries, or large-scale customers [2]. These data are by design for the whole German market area and hence our analysis cannot take into account any spatial differentiation between consumption patterns. Very recent data are subject to revisions, and final data for a given date are only available after ca. 1.5 months.¹ Therefore, at the time of writing, the last available month of final data is December 2022. The publicly available dataset includes the years 2018-2022.

¹THE publishes final data according to ‘M+2M-10WD’, which means that final data for the current month ‘M’ are published two months later (‘+2M’) minus 10 working days (‘-10WD’). This information is provided in a data Excel file available at www.tradinghub.eu/en-gb/Publications/Transparency/Aggregated-consumption-data. Hence, for December 2022, final data have been available since the 15th of February.

SI.1.2. Weather data and other controls

Residential and commercial gas demand is heavily driven by heating demand inducing a high dependence on outside air temperatures and other weather variables. Germany’s National Meteorological Service (DWD) publishes dozens of weather parameters for hundreds of weather stations daily. They are available through an application programming interface (API) that permits downloading specific data with custom programming scripts. We implemented our download routine in Python (see Section SI.5 below). While the model described in the next section could potentially deal with a large number of covariates by means of regularisation methods such as a least absolute shrinkage and selection operator (LASSO), a regularised regression method that constrains the L_1 norm of the coefficient vector helping to select only important regressors, we restrict ourselves to a concise set that accounts for a very large share of the gas demand variation in the control period.

For each day and every weather station, we access the average temperature, as well as the maximum and minimum temperatures. The latter account for extreme temperature changes during a single day. To control for thermal inertia, three lags of average, minimum and maximum temperatures are added to the model² Solar irradiation might be conducive to heating demand reductions not only through its effect on air temperatures. We proxy solar irradiation by the sunshine duration per day in hours. As discussed in the previous subsection, the gas demand data is only available at a national level. Hence, we need to aggregate the covariates spatially. Other studies use population-weighting to average across spatially disaggregated re-analysis data, a blend of historical data points and model outputs [3]. For simplicity, we choose to take the median across weather stations. We prefer the median over a simple average so as to not introduce biases from extreme observations, such as measurements from Germany’s highest mountain *Zugspitze*. Lastly, we include fixed effects for months and weekends as well as national holidays.

²The German building stock equipped with gas-fired heating has varying degrees of insulation. By allowing the model to choose the relative importance of temperatures on preceding days non-parametrically, it accounts flexibly for the average impact of insulation on gas demand.

SI.1.3. Summary statistics

In Table SI.I, we present a few key statistics of our data set. We distinguish between 2018-2021, the business-as-usual period, and 2022, the year subject to behavioural savings. It is evident from the top panel that average gas consumption is somewhat lower in 2022 compared to previous years. For the weather variables in the following panels, the statistics are quite close to each other, indicating good *overlap*, a key requirement for the validity of the method discussed in the next section [4]. For exposition, we plot the mean temperature in the September to December 2022 period against the mean, minimum and maximum temperatures of the same period in 2018-2021 in Figure SI.I

Table SI.I: Selected summary statistics

Variable	Statistic	2018-2021	2022
Gas consumption	avg	1088.44	966.66
	min	165.35	162.93
	max	3273.74	2668.29
	std	764.83	726.95
Mean temperature	avg	10.18	10.61
	min	-9.6	-6
	max	27.2	26.2
	std	7.06	7.03
Minimum temperature	avg	5.63	5.83
	min	-13.7	-10
	max	18.2	16.95
	std	5.99	5.87
Maximum temperature	avg	14.8	15.43
	min	-6	-3
	max	35.5	35.6
	std	8.48	8.46
Sunshine duration	avg	4.96	5.54
	min	0	0
	max	15.23	14.65
	std	4.33	4.38

Figure SI.I: Mean temperatures

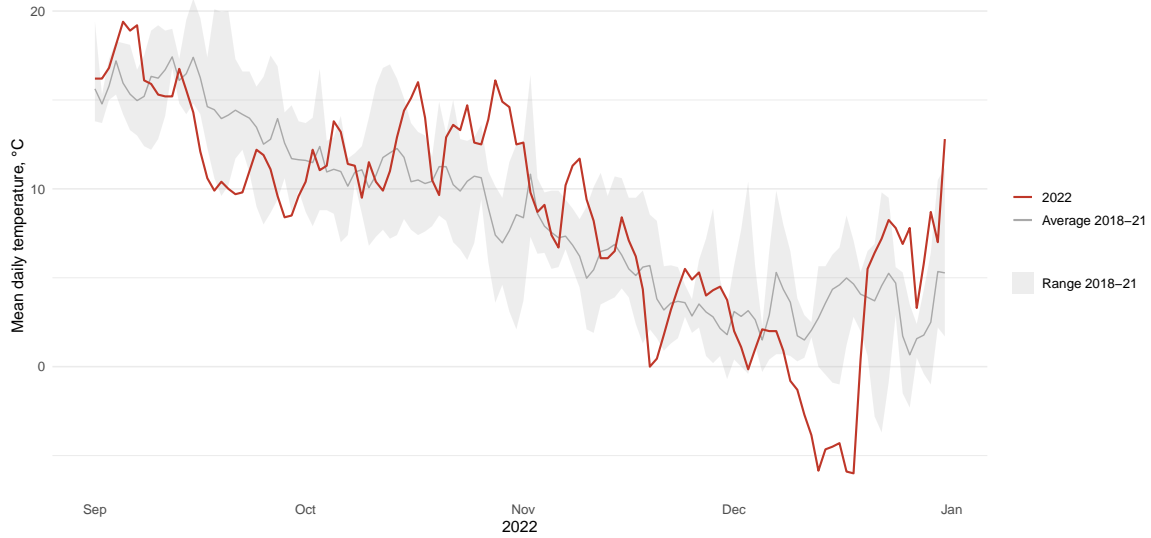


Figure SI.II shows the relationship between gas consumption and the daily mean temperature differentiated by calendar month (colour) and period (marker shape). The figure demonstrates the non-linear relationship between mean temperature and gas consumption.

SI.2. Model description

As evident from Figure SI.II, the relationship between weather variables and gas consumption is non-linear. Traditional methods, such as heating degree day corrections or parametric polynomial models, may introduce biases, especially at the boundary of the support.

We deploy a fully data-driven, non-parametric approach that can not only deal with non-linearities in the relationship between covariates and gas consumption but also with *heterogeneity* in behavioural savings conditional on the covariates, such as temperature or month. Non-parametric models do not require the formulation of a functional relationship between relevant factors, the covariates, and the variable of

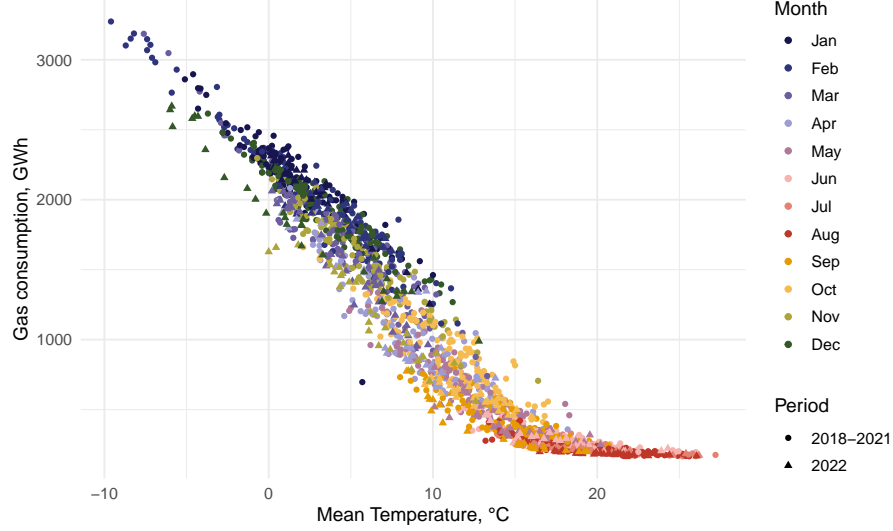


Figure SI.II: Relationship between mean temperature and gas consumption by month and period

interest. Causal forests pioneered by [4], and refined with doubly-robust techniques in [5], extend a classical machine learning method, random forests [6]. We provide short explanations of these terms below.

SI.2.1. Random forests

Random forests predict a variable of interest conditional on a set of covariates by averaging over the predictions of a potentially large number of decision trees. A decision tree splits the data set into subsets, or neighbourhoods, in the covariate domain. In our case, a simple tree could first divide the data set depending on whether a given observation has a mean temperature above or below 10 degrees Celsius. In the below 10 degrees subset, the next split could be based on whether an observation is from a calendar month before March or not. A completely different split could divide the above 10 degrees subset. Further splits may follow. The final subsets, or neighbourhoods, are called *leaves*. For each leaf, the random forest algorithm fits a local model. In the classic implementation, this local model is a simple average of the variable of interest of all observations within this leaf. The more refined version used below fits a local linear model instead. The algorithm

selects splitting rules in order to minimise some prediction error metric, such as the mean squared error. Taken together, the collection of local models can represent complicated non-linear relationships without having to specify a functional form. As shown [6] that the average of a large number of decision trees estimated on bootstrap samples improves the predictive power, a forest usually consists of at least a few hundred decision trees.

SI.2.2. Causal inference and causal forests

Causal forests use random forests to the *prediction* of treatment effects in a potential outcomes framework (e.g. [7]). The *fundamental problem of causal inference* is that we cannot observe what would have happened to a treated unit in the absence of the treatment [8]. In our case, the treatment corresponds to all factors discussed above regarding the looming supply crunch, assuming to start after 23/02/2022. In order to *identify* the treatment effect, i.e. the behavioural savings, modellers have different options ranging from structural models to randomised experiments. Observational studies, like the one at hand, aim to emulate the randomisation of an experiment, e.g. by controlling for all factors that affect the propensity of being treated. Provided we can observe all such factors, the treatment assignment conditional on the covariates becomes as good as random. Two methods (of many) for controlling for the covariates affecting treatment selection are regression and inverse propensity score weighting (IPW). Combining the two leads to the class of *doubly robust* estimators that have the advantage of recovering the treatment effect even if only one of the two methods is correctly specified.

Much like a random forest, a causal forest splits the data set based on rules referring to covariate values. However, the objective sought to optimise by selecting the splits is different. We do not have data on the true treatment effect such that we cannot optimise a prediction error metric. Instead, the causal forest algorithm aims to determine neighbourhoods in the covariate domain in such a way that the estimated treatment effects are as similar as possible within a neighbourhood and as dissimilar as possible between neighbourhoods [9]. The conditional average treatment

effects are especially useful in a context like ours where the magnitude of behavioural savings is expected to vary significantly by weather conditions.

SI.2.3. Model formulation and estimation

Let $Y_t(1)$ be the gas consumption in period t in the presence of behavioural savings and $Y_t(0)$ be the gas consumption in the same period in the absence of behavioural savings. Consequently, the observed consumption can be expressed by:

$$Y_t = W_t Y_t(1) + (1 - W_t) Y_t(0)$$

where $W_t \in \{0, 1\}$ indicates the presence of behavioural savings. In our base case, $W_t = 1$ for all $t \geq 24/02/2022$. We are interested in the effect of behavioural savings conditional on covariates \mathbf{X}_t defined by:

$$\tau(x) = \mathbb{E}[Y_t(1) - Y_t(0) | \mathbf{X}_t = x] \quad (\text{A.1})$$

Yet, $Y_t(1)$ and $Y_t(0)$ are not observable at the same time, such that the function $\tau(x)$ is not directly identifiable.

We assume *strict exogeneity* conditional on the covariates \mathbf{X}_t , i.e. there are no unobserved confounders of W_t and Y_t and after controlling for the covariates the *treatment assignment* is as good as random.

$$\{Y_t(1), Y_t(0)\} \perp W_t | \mathbf{X}_t$$

We further assume that residential and commercial gas consumption Y_t on a day t follows the following partially linear model:

$$Y_t = \tau(\mathbf{X}_t) W_t + f(\mathbf{X}_t) + \varepsilon_t \quad (\text{A.2})$$

The effect of behavioural savings on consumption is measured by a function $\tau(\cdot)$, which may depend on the covariates \mathbf{X}_t . $f(\mathbf{X}_t)$ is a potentially complicated function of the covariates and ε_t is an independently distributed error term. Double

robustness, as discussed above, arises from the following reformulation [5]:

$$Y_t - m(x) = \tau(x)(W_t - e(x)) + \varepsilon_t \quad (\text{A.3})$$

where the regression component is $m(x) = \mathbb{E}[Y_t | \mathbf{X}_t = x] = f(x) + \tau(x)e(x)$ and the propensity score component is $e(x) = \mathbb{E}[W_t | \mathbf{X}_t = x]$. We build the model in two steps:

1. We estimate the nuisance functions $m(x)$ and $e(x)$ using local linear forests. A nuisance function is a function that is not of direct interest for the question at hand but needs to be estimated in order to identify the variable of interest. Local linear forests fit a linear model to the local observations in each leaf. They have proven superior for smooth, non-linear signals [10]. A key feature of our predictions $\hat{m}(x)$ and $\hat{e}(x)$ is the *honesty* property. An honest tree divides the data into two subsamples. The first subsample is used to define the splitting rules and the second subsample is used for the estimation within a leaf. Only with *honest trees* the estimators have the desired asymptotic properties, such as consistency and asymptotic normality required for valid inference [4], e.g. used for the confidence intervals presented in the lower panel of Figure ??.
2. We use a causal forest to find neighbourhoods for the treatment effects. In each neighbourhood, we estimate A.3 where we replace $m(x)$ and $e(x)$ by $\hat{m}(x)$ and $\hat{e}(x)$.³ The formulation has the advantage that we can still recover a good estimate of $\tau(\cdot)$ even if our estimates of the nuisance functions are noisy [9]. We obtain an estimated function $\hat{\tau}(x)$ according to [4, 9]:⁴

$$\hat{\tau}(x) = \frac{\sum_{\{t: \mathbf{X}_t \in \mathcal{N}(x)\}} (W_t - \hat{e}(\mathbf{X}_t))(Y_t - \hat{m}(\mathbf{X}_t))}{\sum_{\{t: \mathbf{X}_t \in \mathcal{N}(x)\}} (W_t - \hat{e}(\mathbf{X}_t))^2}$$

where $\mathcal{N}(x)$ refers to the neighborhood of a particular covariate realisation x

³For notational simplicity, we gloss over the fact that the prediction for t is made on the basis of all observations except for t . See [9] for details.

⁴We assume here that the local linear regression model is solved by ordinary least squares (OLS).



Figure SI.III: $b = 1$ causal decision tree

found by the causal tree.

The causal forest consists of $B = 10^4$ trees in total. For exposition, we show the tree $b = 1$ in Figure SI.III

SI.3. Model fit

Before turning to the robustness checks, we further investigate the model output. The causal forest model discussed in the previous section is trained to predict the *savings effect* induced by the gas crisis. However, as shown in Figure 1 in the main text, we can recover the expected gas consumption in either scenario from the model. We can use these expected gas consumption paths to compare them against the actual consumption in the pre-crisis period (2018-Feb 2022). While the aim of our model is not to maximise the in-sample fit of consumption in the pre-crisis period and it would still produce consistent estimates of the savings effect even if the relevant covariates only accounted for a small share of the observed variation in gas consumption, a good

pre-crisis model fit makes it more unlikely that we omitted a *confounding* factor.⁵ It is, therefore, informative to compare the pre-crisis fit of our model with alternative approaches. Note that in the pre-crisis period, actual consumption is modelled with the ($W_t = 0$) scenario of our model. We can recover the scenario by computing:

$$\hat{\mu}_t(x, 0) = \hat{m}(x) - \hat{\tau}(x)\hat{e}(x)$$

In the left panel of Figure SI.IV, we plot the predictions $\hat{\mu}_t(x, 0)$ against the observed consumption between January 2018 and 23 February 2022. We observe that the fitted values are very close to the black line, which indicates a perfect fit. We compute a root mean squared error (RSME) of 75.2 GWh. In comparison to other methods, such as a simple heating degree day (HDD) correction (middle panel), in which we regress the observed consumption on $\max(15 - T_t^{mean}, 0)$, or the standard load profiles (SLP) used by the network operators (left panel), we note that the causal forest fit is tighter, especially so for the lowest temperatures.⁶

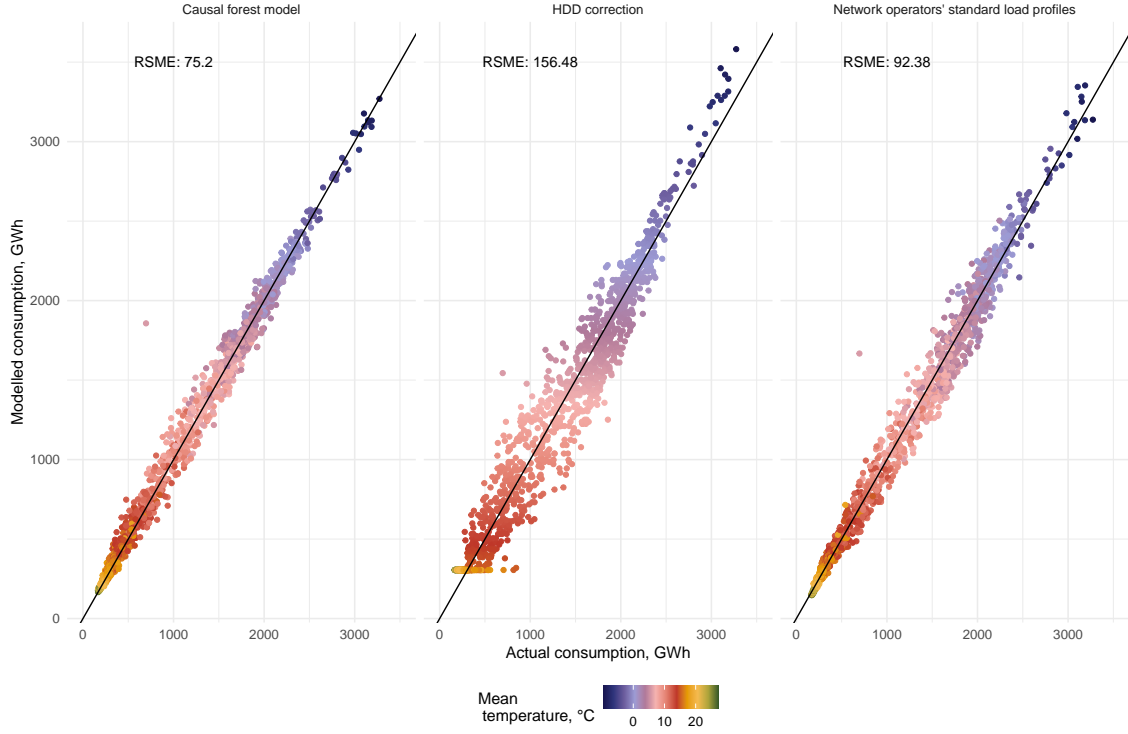
SI.4. Robustness

In contrast to prediction models, where evaluation metrics can be readily calculated on a validation dataset, the quality of the model for $\tau(x)$ cannot be readily evaluated in the same manner as the true values are not available. Following earlier approaches [9], we conduct an auxiliary regression that helps to assess the causal forest fit and test the null hypothesis of no saving effect heterogeneity. We further conduct placebo

⁵A confounding factor is a covariate that affects both the likelihood for a particular day to be one with behavioural savings and the gas consumption. Such a factor, if omitted, induces a spurious behavioural savings effect.

⁶ T_t^{mean} refers to the mean temperature on day t . Like the residual load data, the SLP predictions of the network operators have been retrieved from Trading Hub Europe as well (see Section SI.1). We note that comparing these with our model outputs is not an exact like-for-like comparison as they at least partially rely on (historical) short-term temperature forecasts rather than historical realised temperature data. However, deviations tend to be small and forecasts are subject to repeated validation processes [2]. The SLP predictions result from a set of sigmoid functions with different parameters for different regions and building types and therefore are far more detailed than any sigmoid function-based model we could have built based on publicly available data.

Figure SI.IV: Comparison of causal forest pre-crisis (2018-Feb 2022) fit to other methods



tests (cf. [11]) and run a few sensitivities on the *treatment* start date.

SI.4.1. Omnibus test for causal forest fit

An omnibus test is a test for general model goodness-of-fit evaluation [9]. We fit a simple linear model regressing the estimated left-hand-side of A.3 on the mean predicted savings effect $\bar{\tau} = \sum_t \tau_t$ and the differential effect $\tau_t - \bar{\tau}$:

$$(Y_t - \hat{m}_t) = \alpha_0 \bar{\tau} + \alpha_1 (\tau_t - \bar{\tau}) + \nu_t$$

An α_0 value close to one suggests that the mean effect is correct, while α_1 close to one suggests additionally that the heterogeneity of the effect is well captured. As shown in Table SI.II, both coefficients are fairly close to one and we conclude that

Table SI.II: Calibration test of causal forest with robust standard errors (HC3)

	<i>Dependent variable: $(Y_t - \hat{m}_t)$</i>
	α
Mean forest prediction ($\bar{\tau}$)	1.005*** (0.043)
Differential forest prediction ($\tau_t - \bar{\tau}$)	1.094*** (0.052)

*p<0.1; **p<0.05; ***p<0.01

the causal forest fit is adequate.

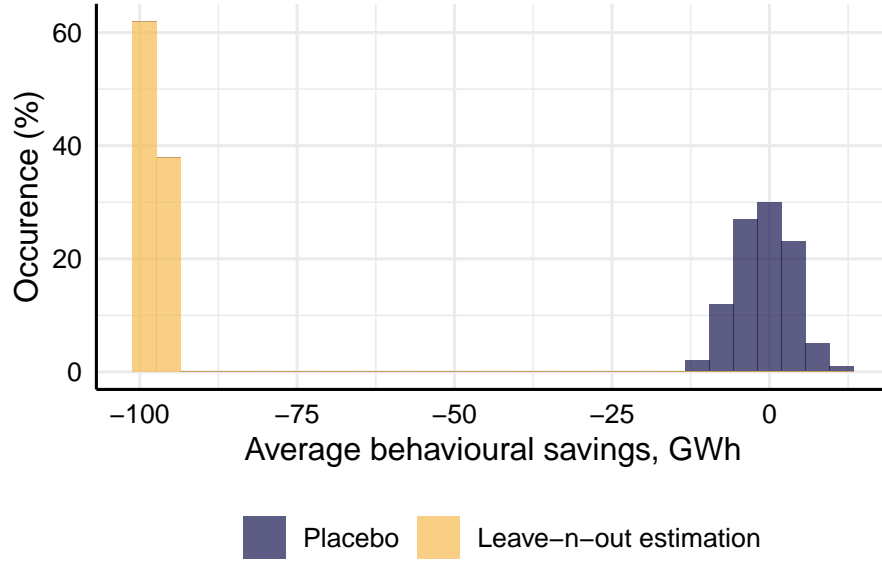
SI.4.2. Placebo testing

Placebo testing is another standard technique in causal inference to test underlying model assumptions [11]. A very basic premise of our model is that the behavioural saving effects do not occur before the start of the energy crisis, the exact start of which is uncertain and subject to additional sensitivities in the next section.

Therefore, we run two sets of auxiliary models. We define $n_{savings}$ as the length of the set of days for which we suppose the presence of behavioural savings $\mathcal{T} = \{t : W_t = 1\}$, hence $n_{savings} = \sum_t W_t$. We define the set of control days as $\mathcal{C} = \{t : W_t = 0\}$. For K times, we take a random sample $\mathcal{S}_k \subset \mathcal{C}$ without replacement. In step (1) of the placebo test, we assign all days $t \in \mathcal{S}_k$ a dummy treatment $W'_t = 1$ and set $W'_t = 0 \forall t \in \mathcal{C} \setminus \mathcal{S}_k$. We estimate our model over \mathcal{C} with W_t replaced by W'_t . We store the average placebo-saving effect. In step (2), we estimate the original model over the set $\mathcal{T} \cup \mathcal{C} \setminus \mathcal{S}_k$. We save the resulting *leave-n-out* average savings effect. We repeat steps (1) and (2) K times to obtain two distributions of average savings effects.

We use the *leave-n-out* estimation instead of our main specification of the model for the true savings effect to make our model comparable to the placebo draws in

Figure SI.V: Placebo test results



terms of statistical power. We set $K = 100$.

Figure SI.V shows the distributions of the Placebo runs in step (1) and the leave-n-out runs in step (2). While all runs in the latter render significant saving effects of around -96 GWh, the distribution of placebo runs is centred around 0 and only 4% of the estimated effects are statistically significant at the 10% level.

We are therefore confident to reject the hypothesis that the estimated savings effects above are just noise.

SI.4.3. Crisis start sensitivities

We have chosen the day of the Russian invasion of Ukraine, 24 February 2022, as the start of our savings period, where $W_t = 1$. However, there are arguments for shifts in either direction. An earlier savings start could be supported by the fact that wholesale gas prices started rising above long-term average levels as early as September 2021 [3]. On the other hand, the need for gas savings for households and commercial sectors only really became evident and a topic in the public domain in the Summer of 2022. Therefore, we test our assumption with respect to the start

date by re-running our model with a monthly sequence of start dates beginning on 24 September 2021 and ending on 24 August 2022. For each iteration, we compute the total cumulative predicted savings in the period from 1 September 2022 until 31 December 2022.

Figure SI.VI: Start date sensitivity

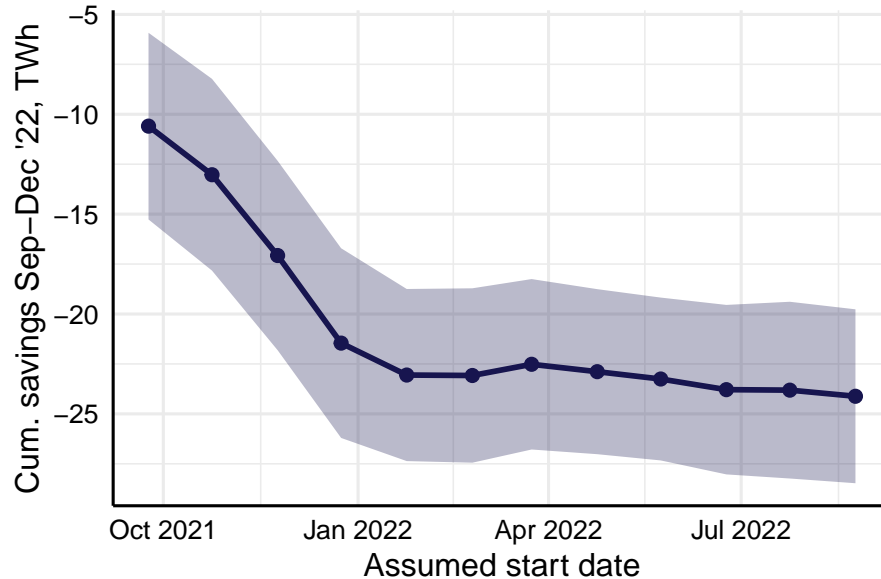


Figure SI.VI shows that the results are very robust to variations of the savings start date in 2022. For start dates in 2021, however, cumulative estimated savings from September 2022 to December 2022 decline rapidly, suggesting that households and commercial sectors did not react to the foreboding developments in wholesale markets at the time.

SI.4.4. COVID-19

The pre-crisis period of our data set includes the COVID-19 pandemic. As households practised social distancing, worked from home, and shops and offices remained closed, the heating behaviour of residential and commercial sectors is likely to have changed compared to the pre-COVID period. Suppose that extended periods of isolation at

home have led to more gas consumption, *ceteris paribus*, even offsetting the reduced demand by commercial buildings. If this were true, our model may deliver biased results as it exaggerates counterfactual gas demand compared to what would have been expected, as the world has gone back to normal in 2022, but for the gas crisis due to the Russian invasion of Ukraine.

Therefore, we conduct a sensitivity test in order to determine if our results hinge on a potential exaggeration of savings due to lockdowns. In Germany, the first lockdown started on 22 March 2020 and ended on 4 May 2020. A second lockdown began with lighter restrictions on 2 November 2020. By January, tighter restrictions were imposed and the lockdown was not lifted before 9 May 2021. Let $\mathcal{L} \subset \mathcal{C}$ be the set of lockdown days in our pre-crisis data set. In the first step, we estimate our model over the set $\mathcal{T} \cup \mathcal{C} \setminus \mathcal{L}$. Further, we define a broader set of pandemic days that comprises all days between 1 March 2020 and 31 December 2021. Let this set be denoted by \mathcal{P} . In a second step, we compute our model estimates over the set $\mathcal{T} \cup \mathcal{C} \setminus \mathcal{P}$.

Table SI.III: COVID-19 sensitivity

Scenario	Est. cum. behavioural savings	Change
Baseline	23.0 TWh	
Excl. lockdown days (\mathcal{L})	22.7 TWh	-1.58%
Excl. all pandemic days (\mathcal{P})	21.5 TWh	-7.09%

As shown in Table SI.III, the effect of excluding the lockdown period is negligible. The effect of excluding the full pandemic period is a bit larger at ca. 7%. However, we do not think it is reasonable to exclude this period entirely. While it is very likely that heating behaviours have changed during the pandemic, it is also probable that at least a part of those changes continues to take effect today, e.g. due to flexible working-from-home policies. We conclude that our model is not substantially biased by the inclusion of the COVID-19 period in the control set \mathcal{C} .

SI.5. Code

We wrote a Python script for gas consumption data downloads and *Deutscher Wetterdienst* API calls. All modelling steps and charting were conducted in R. We make all code available in this repository: gitlab.com/diw-evu/projects/gas-savings.

References

- [1] ACER, Annual Report on the Results of Monitoring the Internal Electricity and Natural Gas Markets in 2021 (Oct. 2022).
URL www.acer.europa.eu/sites/default/files/documents/Publications/MMR_2021_Energy_Retail_Consumer_Protection_Volume.pdf
- [2] BDEW, BDEW/VKU/GEODE Best Practice Guidelines market processes for the management of gas balancing groups - part 1 (Mar. 2021).
URL www.tradinghub.eu/Portals/0/Downloadcenter%20-%20Kooperationsvereinbarung%20und%20Leitf%C3%A4den/KoV%20englisch/KoV_XII_LF_BKM_Gas_Teil1_EN.pdf?ver=iImXwzbftJXgHmEwqSMHgQ%3D%3D
- [3] O. Ruhnau, C. Stiewe, J. Muessel, L. Hirth, Gas demand in times of crisis: energy savings by consumer group in Germany, ZBW - Leibniz Information Centre for Economics Preprint (261082) (2022).
URL <https://ideas.repec.org/p/zbw/esprep/261082.html>
- [4] S. Wager, S. Athey, Estimation and Inference of Heterogeneous Treatment Effects using Random Forests, *Journal of the American Statistical Association* 113 (523) (2018) 1228–1242. doi:10.1080/01621459.2017.1319839.
- [5] S. Athey, J. Tibshirani, S. Wager, Generalized random forests, *The Annals of Statistics* 47 (2) (Apr. 2019). doi:10.1214/18-AOS1709.
- [6] L. Breiman, Random Forests, *Machine Learning* 45 (1) (2001) 5–32. doi:10.1023/A:1010933404324.
- [7] D. B. Rubin, Causal Inference Using Potential Outcomes, *Journal of the American Statistical Association* 100 (469) (2005) 322–331. doi:10.1198/016214504000001880.
- [8] P. W. Holland, Statistics and Causal Inference, *Journal of the American Statistical Association* 81 (396) (1986) 945–960. doi:10.2307/2289064.

- [9] J. Tibshirani, S. Athey, E. Sverdrup, S. Wager, A grf guided tour.
URL www.grf-labs.github.io/grf/articles/grf_guide.html
- [10] R. Friedberg, J. Tibshirani, S. Athey, S. Wager, Local Linear Forests, *Journal of Computational and Graphical Statistics* 30 (2) (2021) 503–517. doi:10.1080/10618600.2020.1831930.
- [11] S. Athey, G. W. Imbens, The State of Applied Econometrics: Causality and Policy Evaluation, *Journal of Economic Perspectives* 31 (2) (2017) 3–32. doi:10.1257/jep.31.2.3.