

## A Preparing metered gas data

This section aims to provide an overview of the process involved in preparing the metered gas usage data for analysis. Metered gas usage data was provided by Gas Networks Ireland. Signed access agreements were obtained from 100 households in our sample who were connected to the gas network<sup>1</sup>. For most dwellings ( $n=65$ ) we have access to 3 years of data (2013-2015), for a subset we could only get access to 2 years of data (2014-2015).

The raw gas features household level interactions between the household and the gas network operator. Each household is identified by their GPRN (Gas Point Registration Number), a unique code for each mains gas meter. The use of GPRN codes also serves to preserve the anonymity of individuals in our data. The raw data features the date of interaction, the read type and the consumption recorded. The read type provides information on whether a customer received an estimated read of consumption, an actual read of consumption or a customer moved dwelling or switched retail provider. Figure 1 provides a graphic representation of the aggregate gas consumption of each household for each day during the sample period.

### A.1 Data cleaning

Cleaning requires a number of steps. First, some observations feature negative consumption values. These values represent corrections of household readings which were inaccurate due to measurement error (typically from overestimates of previous consumption, occasionally from incorrect meter readings). The corrections serve to balance out the consumption for each household and to ensure the accuracy of the aggregate level of consumption for each household. Any observations with extremely large negative readings (less than  $-8,000\text{kWh}$ ) which could not be allocated to previous periods were removed from the sample ( $n = 2$ ).

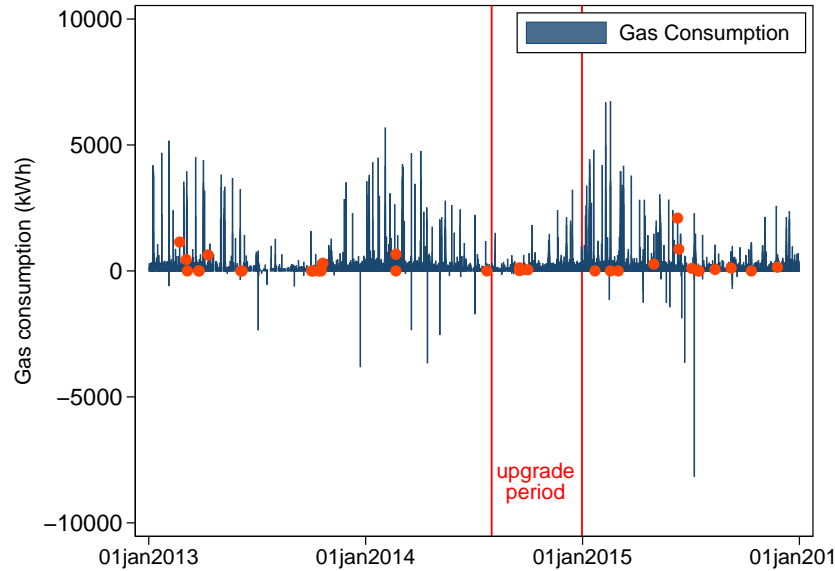
Another potential issue is the concern that some of the residents moved house over the course of our analysis. Based on the gas reading data, we can identify 26 unique households which have changed account holder during the sample period (identified by the red dot in Figure 1 on the date a change was registered). It is also possible that the account holder may have changed between adults in the same household, perhaps to avail of a promotional offer. Of these households, 25 reported in the post-upgrade survey that they are the same respondent as the pre-upgrade survey. Consequently we do not omit these households from our analysis.

At this point, the data are aggregated to a bimonthly billing cycle. Our sample includes both households on post-pay meters and households on pre-pay meters. Post-pay meters generally have a reading every two months,

---

<sup>1</sup>Signed waivers and surveys were collected by Amárach Research.

Figure 1: Daily gas consumption (full sample period)



pre-pay meters are usually read weekly. In order to compare consumption for both groups as a panel, we need to allocate the reading to the period in which consumption occurred, and align them for both groups. An issue we encountered when doing this was a mis-match in billing periods both within and between meter type. For example, the majority of the post-pay meters have a billing cycle that runs bimonthly with January-March being the first period in each year, but for a large minority the cycle starts a month later from February-April. A small number of other households have a different cycle again. To resolve this, we use the median bill period, i.e. the bimonthly cycle beginning in January as our main unit of analysis. All other readings are aligned to this billing cycle based on the date of the reading, and the proportion of consumption that falls into each period. For example, the reading of 17,000 kWh on 03 April 2013 in Table A1 which falls roughly in the middle of period 3 is split between periods 2 and 3 using the ratio of 52:48 as 52% of the days this reading covers relate to period 2 and 48% relate to period 3.

Given our cycle is a weighted average of two period consumption, we have to drop both the beginning and the end period from our sample (periods 1, 7 and 19), for both the two and three year gas time-series.

Prepaid meters reflect consumption that occurred over the previous week. We allocate this consumption to the mid-point of the current and previous reading, then aggregate these by bimonthly cycle to generate comparable gas consumption data over each period for both meter types.

Table A1: Example meter readings and billing cycle

Date	Billing period		
	1	2	3
02 January 2013	500		
08 January 2013	32000		
09 January 2013		9000	
04 February 2013		19000	
01 March 2013		12000	
06 March 2013		32000	
11 March 2013			200
03 April 2013			17000
05 April 2013			4000
08 May 2013			21000

Gas consumption in kWh

At this point any observations with zero gas consumption for any given period are removed ( $n = 3$ ), and a dwelling with a G-rated BER is removed ( $n = 1$ ).

Finally, any remaining negative readings are smoothed across previous periods. In cases where the absolute value of the negative reading exceeds the positive value of the reading preceding it, it is allocated across all preceding periods whose sum is less than the absolute value of the negative reading. For example, the negative reading in period 4 of Table A2 is allocated to periods 2,3 and 4, but not period 1. A three period moving average is also used as a robustness check.

Table A2: Smoothing negative gas consumption values

Period	Cons	Smoothed Cons
1	2364	2364
2	2369	785.66
3	1694	785.66
4	-1706	785.66

The final sample used in the econometric analysis contains 94 households and 1139 observations. Along with the omitted start and end periods a number of dwellings appear to have had meters installed at varying stages within the trial period. These households are included in the analysis but are missing some observations.