Analysis of individual natural gas consumption and price elasticity: Evidence from billing data in Italy

Energy Economics

Analysis of individual natural gas consumption and price elasticity: evidence from billing data in Italy

--Manuscript Draft--

Manuscript Number:	ENEECO-D-22-00085R2
Article Type:	Full Length Article
Section/Category:	Energy Economics
Keywords:	Billing data; intra-day gas demand; Panel data; price elasticity
Corresponding Author:	Luigi Grossi Universita degli Studi di Padova Padova, ITALY
First Author:	Luigi Grossi
Order of Authors:	Luigi Grossi
	Filippo Favero
Abstract:	The price elasticity of residential and non-residential energy demand is crucial for assessing energy policies and consumption forecasts. In this paper, using the monthly billing data of 51,177 end-users in Veneto (Italy), we implemented two panel-data models, analyzing the elasticity between natural gas demand and price variations. The models differ on how the possible endogeneity between price and demand is treated: (1) the total natural gas price is split into several components and we study only the natural gas price part free of endogeneity; (2) instrumental variable methods are used, considering two external instruments to regress the natural gas price from exogenous sources; (3) endogeneity is pictured by a dynamic panel model, correlating consumption with its first lag. We also investigated consumption patterns, underlining the influence of weather factors, market-liberalization, smart-metering technologies and macro-socioeconomic trends. The principal contribution of this paper is twofold: firstly, to supplement the slight empirical literature on natural gas consumption and price elasticity at the micro-level, based on the Italian market, providing a new contribution helpful in defining energy and environmental policies in Italy. Secondly, we provide an empirical method for utilities to exploit information that is often untapped such as billing invoices and technical data. The results show that residential end-users are more reactive to price changes than non-residents. However, the price elasticities found in this analysis are less than 1% for both groups, supporting outcomes found in other countries, i.e., natural gas consumers slightly react to price variations.

Analysis of individual natural gas consumption and price elasticity: evidence from billing data in Italy

Filippo Favero (Università degli Studi di Verona - filippo.favero@univr.it)
Luigi Grossi (Università degli Studi di Padova - luigi.grossi@unipd.it)

October 6, 2022

Acknowledgments

The authors wish to thank the Associate Editor and two anonymous referees for their helpful comments and suggestions. Agsm Aim Energia and V-Reti (Agsm Aim group) are gratefully acknowledged for their assistance, expertise, and for providing the billing and distribution data-sets. However, results and interpretations illustrated in the paper do not reflect the opinion of the Agsm Aim group, but should be uniquely attributed to the authors. Filippo Favero wishes to thank the attendees of a seminar delivered on 8th September 2020 at the Department of Economics, University of Verona, where preliminary results of the manuscript were presented. Luigi Grossi thanks the attendees of the SIS conference, that has taken place in Caserta on June 2022, where extensions of the present paper have been presented and discussed. Some ideas has been used to revise the present manuscript. The research has been partially funded by the Italian Ministry of University (grant FFABR2017). The usual disclaimer applies.

Analysis of individual natural gas consumption and price elasticity: evidence from billing data in Italy

Filippo Favero (Università degli Studi di Verona - filippo.favero@univr.it)
Luigi Grossi (Università degli Studi di Padova - luigi.grossi@unipd.it)

October 6, 2022

Abstract

The price elasticity of residential and non-residential energy demand is crucial for assessing energy policies and consumption forecasts. In this paper, using the monthly billing data of 51,177 end-users in Veneto (Italy), we implemented two panel-data models, analyzing the elasticity between natural gas demand and price variations. The models differ on how the possible endogeneity between price and demand is treated: (1) the total natural gas price is split into several components and we study only the natural gas price part free of endogeneity; (2) instrumental variable methods are used, considering two external instruments to regress the natural gas price from exogenous sources. We also investigated consumption patterns, underlining the influence of weather factors, market-liberalization, smart-metering technologies and macro-socioeconomic trends. The principal contribution of this paper is twofold: firstly, to supplement the slight empirical literature on natural gas consumption and price elasticity at the micro-level, based on the Italian market, providing a new contribution helpful in defining energy and environmental policies in Italy. Secondly, we provide an empirical method for utilities to exploit information that is often untapped such as billing invoices and technical data. The results show that residential end-users are more reactive to price changes than non-residents. However, the price elasticities found in this analysis are less than 1% for both groups, supporting outcomes found in other countries, i.e., natural gas consumers slightly react to price variations.

Keywords: Billing data, price elasticity, intra-day gas demand, panel data.

Research Highlights

- A new dataset based on microdata of monthly gas billings in the Veneto region (Italy) has been created
- Panel data models have been fitted to data to obtain price elasticities of gas consumption
- Weather data, latent factors, smart meter dummy and macro-trend variables have been included in the models
- Sources of endogeneity in the models have been treated using instrumental variables.
- Price elasticities for households and for non-households are consistent with previous results found in the literature

Credit author statement

The authors have shared all the steps of the work. Conceptualization, methodology, empirical analysis, and comments can be equally attributed to the authors. Filippo Favero wrote the R code.

Analysis of individual natural gas consumption and price elasticity: evidence from billing data in Italy

December 5, 2022

Abstract

The price elasticity of residential and non-residential energy demand is crucial for assessing energy policies and consumption forecasts.

In this paper, using the monthly billing data of 51,177 end-users in Veneto (Italy), we implemented two panel-data models, analyzing the elasticity between natural gas demand and price variations. The models differ on how the possible endogeneity between price and demand is treated: (1) the total natural gas price is split into several components and we study only the natural gas price part free of endogeneity; (2) instrumental variable methods are used, considering two external instruments to regress the natural gas price from exogenous sources.

We also investigated consumption patterns, underlining the influence of weather factors, market-liberalization, smart-metering technologies and macro-socioeconomic trends. The principal contribution of this paper is twofold: firstly, to supplement the slight empirical literature on natural gas consumption and price elasticity at the micro-level, based on the Italian market, providing a new contribution helpful in defining energy and environmental policies in Italy. Secondly, we provide an empirical method for utilities to exploit information that is often untapped such as billing invoices and technical data.

The results show that residential end-users are more reactive to price changes than non-residents. However, the price elasticities found in this analysis are less than 1% for both groups, supporting outcomes found in other countries, i.e., natural gas consumers slightly react to price variations.

Keywords: Billing data, price elasticity, monthly natural gas demand, panel data.

1 Introduction

Natural gas at the local level is still the most important fossil fuel used for heating purposes by households. Recent dramatic events related to the Ukrainian war, have raised awareness about the heavy dependence of many European countries on the natural gas supply. Although the ultimate decarbonization requires electric heating that is generated from non-fossil sources, it is quite evident that, in the medium term, gas will be crucial along the path to decarbonization and environmental protection[1]. (Safari et al., 2019) argue that natural gas is a complementary transition fuel towards a zero-carbon system based on renewable sources. They analyse the situation of four countries, two with large gas reserves (Norway and Iran) e two importing gas (India and UK) trough a STEEP analysis and conclude that, ignoring the role played by low emission fuels, the transition to zero emission systems will lead to misleading policy action. However, the relevance of gas in the green transition is controversial. Although natural gas might help the energy transition by reducing emissions compared to coal, according to (Gürsan and de Gooyert, 2021), investments in new infrastructure to bring natural gas to households and firms could lead to long-term implications reducing investments in renewable generation plants and green heating technologies, like heat pumps. This can work against reaching climate goals.

Since 1996, the EU commission has taken measures concerning market access, regulation, consumer protection, and support for interconnection to harmonize and liberalize the EU energy market. These measures aim to foster competition (Opolska, 2017), creating a customer-oriented market with low supply prices. creating a customer-oriented market with low supply prices. Since 2003, the Italian energy market has been liberalized; private sellers can decide to operate in the market, and users can freely decide which supplier to choose. However, the process is likely to terminate only at the end of 2023, when the "Consumer-Protection Market" (hereafter CPM) ceases to exist. Each consumer will have to switch to the "Free Market" (hereafter FM), in which the most convenient supplier can be chosen.

In Italy, the retail energy market is regulated by the Italian Regulatory Authority for Energy, Networks and Environment (ARERA). The principal elements of a typical natural gas bill are sales costs, distribution costs, system expenses, and taxes. The market separation between CPM and FM consumers influences only the selling component.

Italy's natural gas price is subject to high volatility compared to other EU countries.[2] The market transition from CPM to FM will bring further price variations and uncertainty for Italian end-users. This process will affect utility revenues from customers. Therefore, energy operators will need specific forecasting and analysis models to simulate market conditions after the transition from CPM to FM. Moreover, households, industries, and commercial activities may vary their responses according to environmental ideology, expenditure, home characteristics, and energy efficiency programs (Costa-Campi and Trujillo-Baute, 2015). For example, consumers may have different attitudes to the carbon-reduction efforts of authorities. Thus, environmental and economic factors influencing energy use must be carefully defined to construct proper models of natural gas

[1] https://www.roadmap2050.eu/

[2] ARERA Annual Reports, 2021, https://www.arera.it/it/inglese/annual_report/relaz_annuale.

consumption.

The principal purpose of this paper is to carry out a statistical case study from an original data-set based on natural gas billing information for 51,177 final users in a province of Italy in the 2016-2018 time interval. We analyze natural gas consumption patterns, considering several independent variables: consumer characteristics, environmental and socioeconomic factors. In particular, we carefully investigate the elasticity between natural gas prices and consumption both for FM and CPM users. To carry out these analyses, we implemented static panel data models. Furthermore, a dynamic panel model is preliminarily estimated (Appendix B) to investigate the possible influence of past consumption patterns.

Unquestionably, the study of the correlation between demand and prices can have substantial practical importance for utilities, regulators, and policymakers (Otero and Waddams Price, 2001). This study can help energy suppliers in decision-making, contributing to the simulation of future scenarios after the total liberalization of the Italian energy market.

We aim to contribute to the existing literature on natural gas demand in several ways. (1) To the best of our knowledge, we implement the first gas consumption price-elasticity analysis using disaggregated natural gas billing data for Italian consumers; (2) we propose three ways to treat the endogeneity in the relationship between price and demand, a topic often ignored in the literature; (3) the proposed models could easily be replicated in other countries or with different data-sets from utilities that are a source of often untapped data; (4) we test the correlation of gas consumption with various environmental factors (temperature, wind speed, air humidity, rainfall), creating a reference base for subsequent studies; (5) this paper could be the first step in analyzing the effects of smart-metering on consumer sensitivity to prices and consumption; (6) we test the ability to use spatial and temporal macro trends to overcome the absence of individual information such as income or home characteristics.

The paper is organized as follows. Section 2 discusses the existing literature on the estimation of energy demand elasticities in Italy and elsewhere, internationally. Section 3 describes data gathering, starting with the data for, and analysis of, environmental factors, billing and distribution information, and macro trends. Furthermore, it illustrates the construction of the panel data models. Section 4 shows and discuss the empirical results. The paper ends (section 5) with a discussion of the policy, economic, and environmental implications.

2 Literature Review

Discussion of the relationship between energy consumption, consumer characteristics and price variations has been intense in recent research. Energy consumption (e.g., electricity and natural gas) has been investigated in several areas worldwide, at the state and regional level or according to distribution, as well as at the individual customer level. However, unlike in the electricity sector (e.g., Liddle and Huntington, 2021), the variability of natural gas consumption has not been analyzed in detail. Therefore, an increasing amount of research focuses on the natural gas sector, which will play a key role in future decarbonization scenarios, and seems to be more price-response variable than electricity for homes (Trujillo-Baute et al., 2018; Woo et al., 2018).

This section reviews some studies of natural gas consumption to contextualize the empirical case put forward in this paper. Furthermore, we cite some important findings for the electricity sector that can be extended to natural gas. The two sectors are complementary and the two retail markets in Italy have similar characteristics. However, in recent times, gas and electricity in Italy have been partially seen as substitute goods because gas is increasingly being replaced by electricity for heating purposes. This process has been certainly incentivized by recent policy measures subsidizing the green transition in the building industry (Alberini et al., 2011). Several publications have examined the Italian energy sector; however, few studies estimate price elasticities for natural gas consumption, and even fewer use disaggregated data. For example, Besagni and Borgarello (2018) analyzed energy consumption (electricity and heating) on the household scale, using survey data obtained from the Italian National Institute of Statistics (ISTAT). The survey represents the Italian population and focuses on household consumption and expenditure in 2015. They implemented a regression analysis showing that, for electricity expenditure, socioeconomic factors such as income and household structure are more significant than home and appliance characteristics. Conversely, home variables are significant in studies of heating costs. Price elasticity has not been analyzed.

As observed by Bianco et al. (2014a,b), the investigation of energy demand in Italy is lacking although it is fundamental for making informed policy decisions. They analyzed total Italian natural gas consumption, considering a multi-year time series regression both for the residential and industrial sectors. For the residential sector, they considered aggregated data for natural gas consumption at the national level, weather conditions, GDP, and natural gas prices. They developed a demand equation that takes the form of a standard constant elasticity function of consumption with the other variables. The elasticity values were studied, showing that natural gas consumption is primarily influenced by climatic conditions and variations in GDP and is less sensitive to price variations. This may be due to the high costs for users of implementing energy efficiency measures or switching to other forms of energy for heating. The study calculated short-run price-elasticity for the residential and industrial sectors, respectively, -0.15 and -0.34.

Unlike the Italian case, several case studies with energy modeling and concerning price elasticity exist in the international literature (Burke and Yang, 2016). Some countries have been studied deeply, including Brazil (Pereira Uhr et al., 2019), China (Sun and Ouyang, 2016; Liu et al., 2018; Zhang et al., 2018; Zeng et al., 2018; Dong et al., 2019), Greece (Kostakis et al., 2021), Japan (Okajima and Okajima, 2013), Portugal (Silva et al., 2017), Switzerland (Boogen et al., 2017; Filippini and Kumar, 2021), UK (Meier and Rehdanz, 2010) and USA (Fell et al., 2014; Miller and Alberini, 2016).

As evidence of the recent growing interest, some meta-analyses have been completed in the last few years. Labandeira et al. (2017) have recently carried out a meta-analysis in order to summarize the principal empirical results in the literature on energy price elasticities. Their analysis shows that price elasticities from panel data approaches are significantly higher than those from time series but lower than cross-section models. The results for panel data show that, on average, the literature has estimated price elasticity of natural gas demand between -0.2 and -0.5.

Zhu et al. (2018) analyzed household electricity demand and identified the main factors affecting residential energy elasticity. They presented a meta-analysis based on a comprehensive and systematic summary of 103 articles. The results show that price elasticity is sensitive to data types, time intervals, and macro or microdata. Environmental characteristics can also affect short-term price elasticity. Panel data have statistical significance, especially in short-term estimations, and generate higher elasticities than other kinds of data. They found an average short-term electricity price elasticity of -0.245 and -0.582 in the long term for panel data techniques.

Individual modeling is particularly relevant for local-level distribution. It provides more variability than top-down approaches, including inter-individual correlation information, useful to check models and for criticism. Furthermore, once individual estimates are available, both top-down and bottom-up approaches can be applied, forming various attractive aggregates, for instance, sums across groups of customers, sums across time, or both (Brabec et al., 2008).

According to Fumo and Rafe Biswas (2015), the future of residential energy consumption forecasting is the analysis of individual homes, abandoning models that use data from surveys or describe an entire population. This assumption is based mainly on the continuous deployment of smart meters to end-users for natural gas and electricity (Batalla-Bejerano et al., 2016). However, as observed by Ravnik and Hriberšek (2019), individual estimations require a sufficiently large consumer sample size for acceptable results. Indeed, human behavior is a primary factor and can generate non-linear variations in statistical consumption trends. This eventuality can be mitigated in large samples by group behavior.

As observed by Halvorsen and Larsen (2013), price elasticity is influenced by the level of data aggregation. The estimation of price elasticity from macro data, or aggregated microdata, is significantly lower, in absolute terms, than from microdata. In some cases, elasticities for macro data have also proved positive. Aggregation-level dependence is due to homogeneity assumptions about consumer preferences and the distribution of prices in the sample. In light of this, we decided to proceed with a panel data analysis, increasing the sample heterogeneity in prices and consumption, improving the regression quality, and finding stable and coherent price elasticities.

Many recent empirical papers are relevant when panel data models are used to study energy consumption. For example, Labandeira et al. (2012) studied electricity price elasticity at the individual level in Spain, implementing a random effect panel model from monthly billing information. They estimated residential and industrial electricity demand for 453,833 individuals from September 2005 to August 2007. They found a -0.2536 electricity demand elasticity for households and -0.0308 for industrial activities.

Miller and Alberini (2016) studied the aggregation effect on a panel data model with a different data-set at the individual level. The household data-set showed that using more aggregated observations produces a more inelastic demand. However, aggregating the utility level data to the state level makes the demand more elastic.

In another paper, Alberini and Filippini (2011) analyzed residential electricity demand using annual aggregated data at the state level. They implemented static and dynamic data panel models for 48 US states from 1995 to 2007, finding a range of short-run elasticity from -0.08 to -0.15. This is relatively small, compared to the result of Alberini et al. (2011), who analyzed microdata at the household level for the 50 largest metropolitan areas in the US in the period 1997-2007. They regressed energy consumption from environmental and household characteristics as well as price data. From a static model, they estimated own-price elasticity of electricity demand in the -0.860 to -0.667 range, while the own-price elasticity of gas demand is in the -0.693 to -0.566 range.

(McIntyre, 2018) questions the assumption usually made in the literature that the relevant price for estimating the elasticity of energy demand is the contemporaneous price. Using microdata from the UK Living Cost and Food survey in 2007-2014, he finds that estimated price elasticities of gas demand decrease when lagged prices are included in the model.

Another essential issue raised by demand-price analysis is the possible source of endogeneity. As reported by Auffhammer and Rubin (2018), the endogeneity from classical market equilibrium theory can generate simultaneity. Another endogeneity source is typical of retail energy markets, where the price is a demand function in a block tariff scheme. The authors tried to resolve the first kind of endogeneity using natural gas spot prices as instruments. For the second source, they used spot prices and baseline prices to construct a simulated instrument. They found that the natural gas average price elasticity of demand for households ranges from -0.17 to -0.23. Furthermore, they observed that individuals with low and high incomes are both price inelastic during the summer. During the winter, low-income households are more price-sensitive. The endogeneity issue has been also raised by (Alberini et al., 2020) who estimate the price elasticity of natural gas demand in a city of Ukraine. Using billing data information collection by an ad-hoc survey the authors estimate the elasticity of gas demand to prices by a panel regression. The source of endogeneity is related to the block pricing system applied in Ukraine as the consumer selects both its consumption level and the corresponding tariff. The endogeneity issue is addressed by a two stage least square estimation where the instruments used in the first stage are allowance and discount off the regular tariff obtained by a group of households.

The research on price elasticity might have practical policy implications. For example, Xiang and Lawley (2019) estimate the effects of the 2008 carbon tax in the British Columbia on residential natural gas consumption per capita. Implementing a panel data analysis, they used several input variables with data from 1990 to 2014, such as gas demand, weather conditions, prices, building and population characteristics. Their results are consistent with the supposition that the residential demand is price and tax inelastic. Indeed, factors such as weather and household characteristics are essential in the short and medium-term, compared to gas and household incomes. However, they were able to assess the long-term impacts of a carbon tax. The overall results suggest that a carbon tax of 1 cent/m^3 reduces per capita natural gas consumption by 3%. Hence, the result can be interpreted as negative price-demand elasticity. A recent study by (Liu et al., 2022) has found that a carbon tax to stimulate the use of “clean” fuels in China, using macro-panel data in the time span 2005-2019 has implied an increasing substitution of natural gas for coal and LPG over time; on the other hand electricity has gradually replaced the use of natural gas.

Although not focused on price elasticity, an important contribution is given by (Auffhammer, 2022) who estimates the effect of climate change in California on electricity and gas consumption. Using a massive micro-dataset containing information from gas and electricity bills paid by households, the author finds that the increased demand for cooling systems due to high temperature will induce a higher demand for electricity, but projections show that this increased electricity consumption will be more than offset by savings in directly consumed natural gas.

Some of the principal results from natural gas consumption analysis and price elasticity are summarized in Table 1.

3 Data and Methodology

We collected monthly natural gas distribution information from about 1,800,000 monthly bills from the Vicenza and Treviso districts in the Veneto region of Italy. Agsm Aim Energia Srl and V-Reti Spa - respectively the principal energy provider and local energy distributor in the area - provided the data-sets.

The population of the area is about 250,000, and the total number of natural gas end-users was approximately 140,000 on December 31, 2018. The balanced sample size analyzed in this paper is about 51,177 consumers, considering only consumers with constant combination of User-Home-Seller.

The monthly data variation is from January 1, 2016 to December 31, 2018 (36 time frames). To the billing and distribution information, we added environmental factors collected from seven weather stations, and macro socio-economic trends from Italian institutional websites.

Two kinds of consumers can be identified in the data-set: "Households" and "Non-households". The household cluster includes end-users that consume natural gas for residential and private use (e.g., heating systems, cooking, cooling, etc.). The non-household cluster includes commercial consumers (e.g., industry, shops, restaurants, artisans, service activities, etc.).

The variables used in this paper and a summary of statistics are set out in tables 2 and 3. A detailed description follows in the next subsection.

3.1 Data Collection

Billing data

Natural gas consumption (Q) is the quantity in Sm^3 (standard cubic meters) of gas consumed monthly by each end-user. We implement the panel data models utilizing logarithmic natural gas consumption as the dependent variable to directly study gas price elasticity (1% variation in natural gas prices would lead to a percentage variation in natural gas consumption). Moreover, it is helpful to mitigate the effects of outliers and standardize the sample toward a normal distribution.

Individual prices are directly derived from the billing data-set. The data are heterogeneous for individuals (households, industrial, and commercial activities) and market conditions (CPM and FM). Moreover, end-users joined the FM at different times with different tariff structures. For price-demand elasticity the total price (*price.tot*) faced by each individual is the relevant datum - i.e., the sum of the selling price (*price.sel*) and distribution price (*price.dis*) -.

The distribution price is divided into several brackets in a block tariff system. The allocated consumption in each group is consecutive, starting at the beginning of the calendar year, so consumption in the first months of the year is in the first group, in later months in subsequent groups and so on.

Figure 1 shows box-plots describing variations in the price components over time. The sample is divided into CPM (1.a and 1.c) and FM consumers (1.b and 1.d).

Author(s)	Data	Technique	Price elasticity
Natural Gas consumption			
Meier and Rehdanz (2010)	Residential energy consumption data in UK Survey 1991-2005 5,000 households	Random effect panel data	-0.34 to -0.56
Alberini et al. (2011)	US household-level energy consumption Survey 1997-2007 98,772 observations in 54 cities	Static and dynamic panel model Variables: energy prices and home characteristics	$\simeq -0.7$ (short-run) $\simeq -0.5$ (long-run)
Auffhammer and Rubin (2018)	Residential natural gas data in California Billing data 2003-2014 2,526,503 households	Instrumental panel data with different instruments	-0.23 to -0.17
McIntyre (2018)	Residential natural gas and electricity data in UK UK Living Cost and Food survey 2007-2014	Quadratic Almost Ideal Demand System (QUAIDS) model	-0.66 to -0.51
Zhang et al. (2018)	Short and long run price elasticities of gas consumption in China for various sectors China Energy Statistics Yearbook 1992 - 2015 Sectoral macro data on gas consumption	Gas demand function estimated by the ARDL approach	Short and long run price elasticities of residential sector: -0.233 (same value)
Zeng et al. (2018)	Chinese residential energy consumption Survey 2014 Cross-sectional 1,035 observations	Logarithmic OLS, IV	-0.898 to -0.965
Dong et al. (2019)	Price elasticities of gas consumption in 30 Chinese provinces China Energy Statistical Yearbook 1999 - 2015 Unbalanced macro panel dataset	Panel common correlated effects mean group estimator	0.02 to 0.04
Alberini et al. (2020)	Residential natural gas data in the city of Uzhhorod (Ukraine) Billing data and Survey data 2013-2017 500 households	Instrumental panel data with instruments	-0.22 to -0.07
Filippini and Kumar (2021)	Residential natural gas consumption in Switzerland Household level data 2010 - 2014 Unbalanced panel of 958 households	Fixed Effect model and Random Effect model with Mundlak's adjustment	$\simeq -0.73$
Kostakis et al. (2021)	Residential gas consumption in Greece Household level panel data 2012 - 2019 185 cohorts of household's heads	Static and dynamic panel models	-0.65 to -0.54
Auffhammer (2022)	Residential natural gas data in California Billing data 2004-2015 More than 90 millions bills	Two-stage log-linear regression models estimated locally	-

Table 1: Principal results for energy price elasticity and gas demand variation from micro- and macro-data.

Description	Source	Variable	Unit
Dependent variable			
Natural gas consumption	Billing data	Q	$[Sm^3]$
Price			
Natural gas price (total)	Billing data	*price.tot*	$[€/Sm^3]$
Weather variations			
Heat Degree Days	ARPAV	*hdd*	$[°C]$
Humidity average	ARPAV	*hum*	[%]
Rain precipitation	ARPAV	*rain*	$[mm]$
Wind velocity average	ARPAV	*wind*	$[m/s]$
End-user characteristics			
Presence of a smart-meter	Distribution data	*smart.meter*	dummy
Consumption self reading by end-user	Distribution data	*self.read*	dummy
Number of PODs in the same building	Distribution data	*n.dwell*	set of dummies
Social welfare aid program participation	Billing data	*welf.aid*	[dummy]
Free-market end-user	Billing data	*free.market*	dummy
Socio-economic trends			
Average size of home	ISTAT	*av.s.dwell*	$[m^2]$
Building-age index	ISTAT	*build.age.index*	$[-]$
Gross average income	Italian ministry of finance	*gross.income*	$[€ \cdot 10^{-3}]$
Population dependence index	ISTAT	*dep.index*	$[-]$
Time fixed effect			
Quarter	–	*q1, q2, q3, q4*	dummy

Table 2: Variables used in panel data models for the individual i and month t.

	Households					**Non-households**				
	mean	max	min	St. Dev.	St. Dev./mean	mean	max	min	St. Dev.	St. Dev./mean
Q	89.95	997.45	2	96.5	1.07	400.24	67,793	2	1620.69	4.05
$ln(Q)$	3.91	6.91	0.69	1.18	0.3	4.45	11.12	0.69	1.6	0.36
price.sel	0.26	0.38	0.18	0.03	0.12	0.29	0.38	0.17	0.04	0.14
price.dis	0.13	0.17	0.05	0.03	0.23	0.13	0.17	0.04	0.03	0.23
price.tot	0.39	0.55	0.25	0.04	0.1	0.42	0.56	0.24	0.05	0.12
hdd	180.79	567	0	178.72	0.99	180.59	566.8	0	178.42	0.99
wind	0.86	1.8	0.3	0.31	0.36	0.86	1.8	0.3	0.31	0.36
rain	83.26	271	0	53.9	0.65	83.85	271	0	53.91	0.64
hum	23.81	51	6	7.24	0.3	23.77	51	6	7.2	0.3

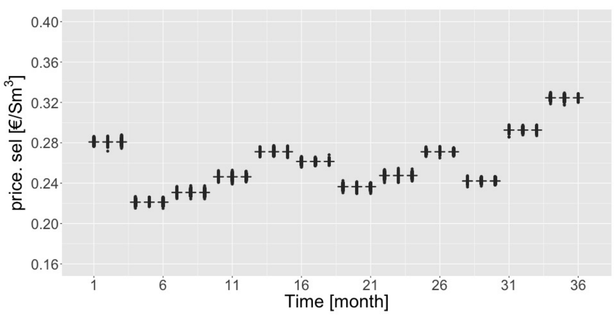
Table 3: Summary statistics for households and non-households.

The cross-section variation of *price.sel* is almost negligible for CPM consumers. The small variation is due to the sample spatial variability that slightly influences the price for different local administrations. Conversely, the *price.sel* for FM users derives from a direct agreement between the seller and consumers, with greater variability. In both cases, the quarterly structure of the sales tariff is clear, i.e., the price remains the same for three-month periods. In figure 1.b, the mean of the selling prices is slightly different for CPM and FM groups. FM sellers typically encourage consumers to switch with affordable prices, similar or equal to CPM.

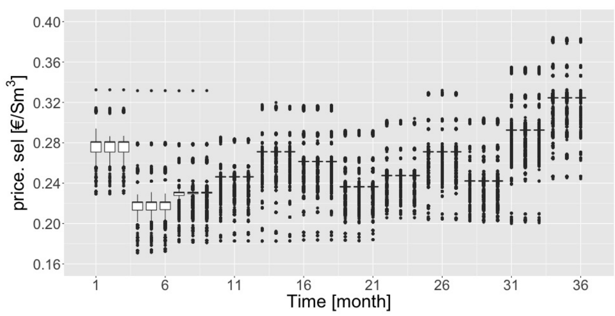
The *price.dis* component does not depend on the consumer choice, either CPM or FM. Figures 1.c and 1.d show how the block tariff system for the distribution prices works. There is a correlation between distribution prices and cumulative natural gas consumption during the year. In January (months 1, 13, 25 in the sample), the *price.dis* is conspicuously lower. End-users are typically in the first consumption step. At the end of the year (October-December), the price flattens around a steady-state value reached by almost all users, corresponding to a quantity of gas in the range from 1000 to 2000 $Sm^3/year$.[3]

The free-market variable ($free.market$) is a dummy variable indicating the market characterization of each end-users. The value of consumers in the free-market increased over the months, especially for the household group, reaching the 40% at the end of 2018. The non-household group is almost static over time and was in the free-market since before the period under analysis. In January 2022, all Italian energy consumers will be in the free market.

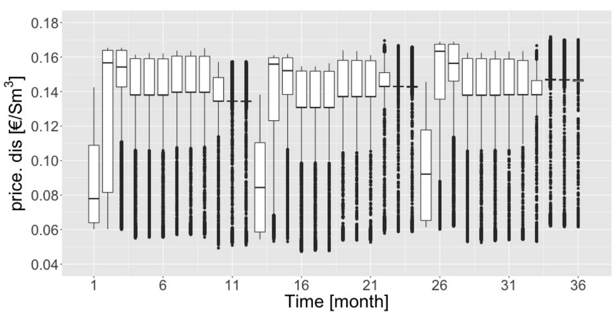
The social welfare variable ($welf.aid$) is considered as a low-income index constant over time. Indeed, the balanced data-set is constructed for static end-users; hence, it is reasonable to consider access to social welfare subsidies as a threshold for low-income end-users during the entire period under analysis. Of the 49,156 residential consumers in the sample, 2,165 end-users received a social welfare subsidy for energy consumption.



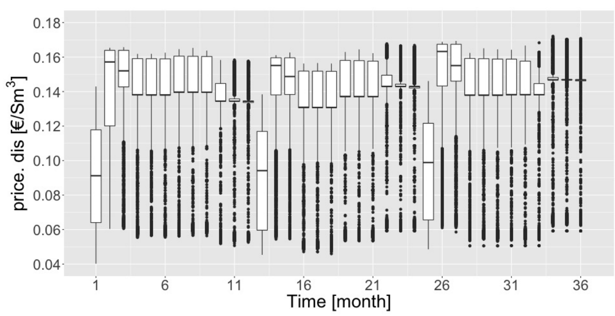
(a) CPM - *price.sel*



(b) FM - *price.sel*



(c) CPM - *price.dis*



(d) FM - *price.dis*

Figure 1: Box-plots of the price variation over time for CPM and FM users.

Weather data

The territories analyzed in this paper are located in the Po Valley. The climate is continental with cold and humid winters and hot, muggy summers. The hills and mountains surrounding the area create a natural barrier to rain and storms.

The environmental data-set was collected from the ARPAV (Provincial Agency for Environmental

[3]ARERA natural gas prices. `https://www.arera.it/it/prezzi.htm.`

Prevention and Protection in Veneto).[4]

Generally, in energy demand analysis, the most significant environmental variable is temperature (Swan and Ugursal, 2009). This is true especially in the natural gas sector since it is mainly used for residential, commercial, and industrial heating. Gas consumption comprises two parts: one small and unrelated to temperature (e.g., cooking, hot water, industrial production) and another, larger, related to the outside temperature (e.g., heating purposes).

The variable used to describe temperature variations is the "heating degree days" (hdd), calculated as the difference between the daily mean temperature and a base temperature of $18°C$ (see equation 1).

$$hdd = \begin{cases} \sum_{d=1}^{31}(18 - T_d) \quad if \;\; T_d < 18 \\ 0 \quad otherwise \end{cases} \tag{1}$$

The temperature is not the only environmental variable with a possible key role; we also collected the monthly mean of humidity (hum), wind speed ($wind$), and the quantity of rainfall ($rain$). As shown in previous studies (Soldo, 2012) (Valor et al., 2001), hdd is the main driver of energy consumption; however other meteorological variables may improve estimation results when energy consumption is analyzed.

Distribution data

The billing data were interpolated with the distribution information provided by the local distributor. The distribution data include: consumer spatial location (address and local administration), presence of smart meters, self-meter reading by end-users, and manufacturing, which is significant for non-household consumers.

The local administration is identified to assign the relevant weather station and the specific macro trends to each end-user. Furthermore, the address is helpful in constructing the $n.dwell$ variable, representing the number of homes in the same building. The $n.dwell$ variable is constructed using the spatial indication of addresses, street numbers, and ZIP codes.

The correlation between the $n.dwell$ and gas consumption is supposed to be negative for two reasons. First of all, because single homes usually have more surface area to heat. Secondly, several homes in the same building contribute to overall heating, decreasing heat loss.

The presence of smart meters is indicated with a dummy variable ($smart.meter$). Energy smart-meters will become essential in the future, due to the development of smart city technologies ((Batalla-Bejerano et al., 2020)). These innovative devices enable two-way communications with real-time control of different variables (flow, volume, temperature, pressure), recorded every 15 minutes. In Italy, ARERA has required natural gas providers to replace at least 85% of traditional meters with smart ones by 2023. Therefore, Italian utilities have run several replacement campaigns. For example, since 2016, the local distributor of the analyzed area has started a replacement campaign aiming to dismiss most of traditional meters. Therefore, this technology is not uniquely reserved to new buildings, but is accessible to all users. As shown in figure 2, the number

[4] Regional Agency for Environmental Prevention and Protection of Veneto, `https://meteo.arpa.veneto.it`. Last accessed: May 2022.

of end-users with smart metering has sharply increased over 2016-2018.

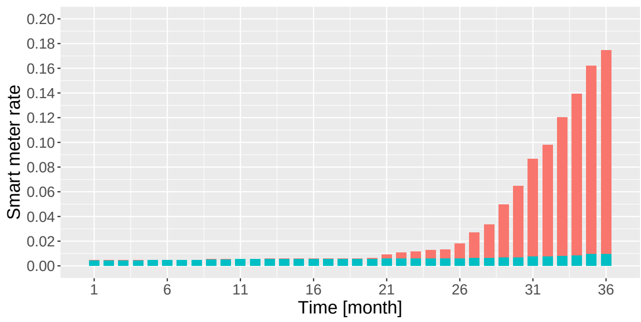


Figure 2: Monthly smart meter rate in the area (2016-2018).

The self-meter reading is indicated by a dummy variable ($self.read$), based on bimonthly observations. The advantage for consumers is that they avoid estimated invoices, which must be adjusted after the official meter reading. Furthermore, sellers usually provide a discount for each self-reading. The issue will be less relevant with the spread of smart meters and real consumption billing without the need for supplementary reading by users or suppliers.

The classification of manufacturing activities ($prod$) is relevant for the non-household cluster, where it is a dummy indicating whether a given activity uses natural gas for production purposes. The information is obtained from consumers' declarations in the database of the provider. The gas sample used in the following models includes 650 users who declared that the natural gas is used both for heating and production purposes. Hence, a linear regression can be expected to highlight a significant increase in consumption for users with this particular feature.

Macro trends

In the sample, there are differences between consumers in heating area, GDP, quality of construction, and demographics. Therefore, we studied some relevant macro trends for each specific local administration. These data were applied to the regression models, to partially adjust the lack of individual information in the data-set. As macro trends, we adopt average per capita gross income, average size of the home, the building-age index, and the population dependence index. Details of annual average gross income ($gross.income$) for each ZIP code were collected from the Italian Ministry of Finance website.[5] There are two possible benefits from using the average income parameter in the panel data regression. First, the parameter may be useful to mitigate the absence of income information, which is only partially satisfied by social welfare aid ($welf.aid$). Second, it can be interpreted as the variation in general economic welfare in the areas of interest over time and space. Therefore, we expect to find a positive correlation between natural gas demand and average income. Indeed, with a positive variation in economic well-being, an increase in consumption is desirable.

Individual information on the size of the home is a fundamental parameter for natural gas consumption, but unfortunately, it was not available in the billing and distribution data. We therefore decided to use the average surface area parameter ($av.s.dwell$) to mitigate the absence of individual

[5] Italian Ministry of Finance. Italian gross income pro-capite data per ZIP code. https://www1.finanze.gov.it/. Last accessed: September 2022.

information. Moreover, the parameter can be useful to support the *n.dwell* variable in approximating the home heterogeneity over the sample. The *av.s.dwell* is constant over time for each ZIP code, and it was obtained from the 2011 ISTAT survey.[6]

The building-age index (*build.age.index*) represents the building renovation ratio. Like the size of the home, it is a fundamental parameter for natural gas consumption. Recently renovated buildings are expected to require significantly lower natural gas for heating. The index was calculated from the 2011 ISTAT database.[7] It is the ratio between the number of buildings less than 20 years old and the number of buildings over 20 years old in 2011. It was determined for each ZIP code, and was constant over time.

The population dependence index (*dep.index*) represents the social and economic burden of the non-active population (0-14 years of age and 65 and over) on the active population (15-64 years of age). For example, in 2017 in Vicenza, there were 57.26 dependent individuals for every 100 inwork. The data vary annually and spatially over the ZIP codes. The data were collected from the Open-Data archive of the Veneto region.[8]

3.2 Model Specification

This section describes the implementation of the empirical models adopted to study natural gas consumption patterns and the price elasticity of demand.

Relying on the existing literature, we implement the models in the double logarithmic form for the price, directly studying the elasticity, and in a log-linear form for the other variables. The models are based on the assumption that energy consumption, for individuals, is principally influenced by four factors (see table 2): price variations (p), weather variations (w), end-user characteristics (d), and social-economics factors (e). Therefore, the basic equation that drives the proposed models is:

$$\ln(Q_{i,t}) = \alpha \ln(p_{i,t}) + \beta w_{r,t} + \gamma d_{i,t} + \phi e_{r,t} + \lambda q_t + c_i + u_{i,t} \tag{2}$$

where α, β, γ, and ϕ are the matrices of the regression parameters and c_i and $u_{i,t}$ are the individual unobserved effect and the idiosyncratic error, respectively. We also included a quarter dummy q_t, which indicates whether the observations are in the 1st, 2nd, 3rd, or 4th quarter of the year. The subscripts i, r, and t indicate individual, regional (characteristic of each ZIP code), and temporal variability, respectively.

When dealing with panel models, serial correlation and heteroskedasticity in the idiosyncratic error terms are problems that must be considered. The results of the Breusch Pagan (Breusch and Pagan, 1979) and Wooldrich tests (Wooldridge, 2002) suggest that the sample is affected by both, then we use cluster-robust covariance matrix estimators.

Another important issue that can affect estimation quality is endogeneity misidentification. Many papers have studied energy demand-price elasticity; however, only a few have dealt with this question, providing potential bias. We treat the topic explicitly, trying to overcome the problem in the

[6] Censimento della popolazione e delle abitazioni, 2011,`https://www.istat.it/it/censimenti-permanenti/censimenti-` Last accesses: May 2022.

[7] ISTAT, Censimento della popolazione e delle abitazioni 2011

[8] Regione Veneto, Open data Veneto, `https://dati.veneto.it`.

models illustrated below.
In particular, starting from equation 2, three sources of potential endogeneity might affect the estimation:

1. The first concerns the natural gas price, which can suffer from simultaneity in the classical theory of price-demand equilibrium. The error term variation can shift the demand curve, inducing a variation in the equilibrium price.
   This kind of endogeneity is the weakest form in this empirical research for two reasons. First, ARERA settles the price structure quarterly in Italy before the start of the period itself. The possible bias in the estimation can be induced by the projection of consumption presumed by ARERA in the price decision process.
   Second, we analyze micro panel data, where single consumers cannot shift the price curve constructed for northern Italy. Therefore, the influence of our sample demand can be considered negligible in price construction, especially for CPM consumers. The endogeneity problem could arise significantly in the FM group; indeed, the selling price is settled by sellers and depends on several factors such as the type of end-user and the market price. However, as seen above, prices for CPM and FM differ slightly in the period 2016-2018.

2. The second source is due to the specific structure of distribution and system charges, which are direct functions of demand.
   The price as a demand function is the most dominant form of endogeneity affecting price elasticity with an upward bias. This type of endogeneity is dealt with in all the models illustrated in the following subsections.

3. The third source concerns the omitted-variable problem, which occurs when relevant variables correlated with the other explanatory variables are omitted. In this case, the error terms are correlated with the dependent and explanatory variables, and the regression becomes biased and inconsistent.
   The easiest way to handle the problem is to use the fixed-effect method, eliminating the bias when the omitted variables are constant over time. However, some relevant non-time constant variables could have been omitted (e.g., household income, industrial production, number of appliances, etc.). To overcome, or at least mitigate, this lack of basic individual information, we utilize the information as macro trends instead of individual variables. Moreover, quarter-indicator dummies are used in the models. These dummies are essential to capture different individual behavior over the year. For example, heating systems are usually active only in the 1st and 4th quarters. 3rd quarter consumption is low due to the summer both for households and non-households. Furthermore, Italy's price structure is established quarterly so a quarter dummy is required in order to capture some in-year quarter seasonality that may affect price elasticities.

Studying the model of equation 2, without accounting for endogeneity issues, the price-demand elasticity results to be positive[9], suggesting that the second and third types of endogeneity (the

[9]Respectevely 0.428 and 0.55 for fixed effect (within estimator) and random effect methods (Swamy and Arora estimator).

price function of consumption and omitted variables) are primary problems.
The following subsection sets out a theoretical description and the estimation results of the models. The sample is clustered into two groups: "Households" and "Non-households". The models are built following the indication of the textbooks of Wooldrich (Wooldridge, 2002), and Baltagi (Baltagi, 2005). Data processing is carried out with R statistical software following Croissant and Milo (Croissant and Millo, 2018).
To improve the interpretation of results presented in the next section, variables related to weather, characteristics of houses, and macro trends (except dummies) have been standardized.

4 Results and Discussion

4.1 Model 1

The easiest way to overcome the endogeneity problem is to eliminate the natural gas distribution price component (*price.dis*), which is a direct function of consumption, and retain only the selling price component (*price.sel*). Indeed, *price.sel* may be slightly affected by the simultaneity problem, although it is not a demand function.
Simultaneity is considered a marginal problem both for CPM and FM consumers. For the CPM, the sales price is established a priori quarterly by ARERA. The *price.sel* is calculated based on spot prices and other macroeconomic trends. Therefore, the influence of individual demand on price is considered negligible. For the FM, the *price.sel* is the result of a direct agreement between consumers and the utility. However, FM and CPM prices slightly differ.
To sum up, in *Model 1*, we accept the bias introduced by the weak simultaneity problem between the *price.sel* and natural gas quantity. Nevertheless, we eliminate the reverse selection problem derived from *price.dis* that is omitted from the model.
To compare price elasticities from different methods, we estimate the model with the two classical panel model techniques that account for individual unobserved effects: fixed effects (FE) and random effects (RE). Table 4 shows the results of *Model 1* for individuals classified as households and non-households. Price elasticity is represented by the coefficient of the component *price.sel*.
The results of table 4 for the household group indicate that price elasticity is in the low range observed in the existing literature. However, we are unable to observe the elasticity of the distribution in the total natural gas price. Hence the elasticity is lower in magnitude than expected.
In the non-household group, the results deviate downward compared to the household group, confirming past outcomes in the literature, i.e., non-residential users result to be less responsive to price variations than residential users.

4.2 Model 2

Model 2 explicitly resolves the endogeneity problem using the instrumental variables method that achieves consistent estimations when there is a correlation between the idiosyncratic error and the explanatory variables. In particular, the regressor *price.tot* is mainly affected by the second type of endogeneity, a demand function for the *price.dis* component. The within estimator can wipe out

	Model 1			
	Households		Non-households	
	FE (1)	RE (2)	FE (3)	RE (4)
ln(price.sel)	-0.3393^{***}	-0.3206^{***}	-0.1761^{***}	-0.2233^{***}
	(0.0055)	(0.0055)	(0.0352)	(0.0343)
hdd	0.4879^{***}	0.4883^{***}	0.4610^{***}	0.4609^{***}
	(0.0015)	(0.0015)	(0.0086)	(0.0085)
hum	-0.0199^{***}	-0.0193^{***}	-0.0055	-0.0053
	(0.0005)	(0.0005)	(0.0030)	(0.0030)
rain	-0.0014^{*}	-0.0020^{***}	-0.0114^{**}	-0.0130^{***}
	(0.0006)	(0.0006)	(0.0035)	(0.0035)
wind	-0.0106^{***}	-0.0120^{***}	-0.0091	-0.0160
	(0.0012)	(0.0012)	(0.0087)	(0.0087)
winter	0.0044	0.0044	0.1170^{***}	0.1196^{***}
	(0.0028)	(0.0028)	(0.0164)	(0.0164)
rain:winter	0.0038^{***}	0.0038^{***}	0.0046	0.0063
	(0.0008)	(0.0008)	(0.0050)	(0.0050)
winter:wind	0.0685^{***}	0.0698^{***}	0.0651^{***}	0.0708^{***}
	(0.0020)	(0.0019)	(0.0150)	(0.0149)
n.dwell		-0.1910^{***}		-0.1558^{***}
		(0.0040)		(0.0330)
smart.meter	-0.1989^{***}	-0.1980^{***}	-0.1726^{***}	-0.0150
	(0.0046)	(0.0045)	(0.0315)	(0.0317)
self.read	-0.0527^{***}	-0.0537^{***}	-0.0055	0.4264^{***}
	(0.0046)	(0.0045)	(0.0652)	(0.0946)
av.s.dwell		0.0428^{***}		0.0055
		(0.0038)		(0.0350)
build.age.index		-0.0975^{***}		-0.1008^{*}
		(0.0051)		(0.0403)
welf.aid		-0.0280		
		(0.0150)		
prod				1.2687^{***}
				(0.0635)
free.market	0.0297^{***}	0.0335^{***}	-0.1258	0.1195
	(0.0037)	(0.0034)	(0.0842)	(0.0993)
gross.income	0.0533^{***}	0.0341^{***}	0.0568^{***}	0.0100
	(0.0022)	(0.0020)	(0.0138)	(0.0129)
dep.index	-0.1711^{***}	-0.1380^{***}	-0.1759^{***}	-0.1926^{***}
	(0.0066)	(0.0048)	(0.0444)	(0.0354)
q1	0.0034	0.0050^{**}	0.0490^{***}	0.0572^{***}
	(0.0018)	(0.0018)	(0.0092)	(0.0092)
q2	-0.2788^{***}	-0.2735^{***}	-0.2373^{***}	-0.2262^{***}
	(0.0028)	(0.0028)	(0.0199)	(0.0198)
q3	-0.5086^{***}	-0.5056^{***}	-0.4174^{***}	-0.4124^{***}
	(0.0030)	(0.0030)	(0.0186)	(0.0186)
R^2	0.5603	0.5537	0.4778	0.4718
Adj. R^2	0.5477		0.4627	
Num. obs.	1,769,616	1,769,616	72,756	72,756
Hausman test	400.39 ($p < 0.001$)		802.3 ($p < 0.001$)	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 4: Model 1 results for household and non-household groups. Dependent variable: logarithmic monthly natural gas consumption ($ln(Q)$). Clustered standard errors in ().

the individual effect c_i; however, the correlation between the idiosyncratic error and the explanatory variables remains. Therefore, the fixed effect estimator in table 4 is biased and inconsistent.
The problem can be solved by breaking the endogenous link between individual price and demand, instrumenting the natural gas price. The standard approach is to specify a set of instruments correlated with the independent variables but uncorrelated with the idiosyncratic errors. To make the estimation, we apply the $2SLS$ technique.
We applied the $2SLS$ technique in the panel models via two different estimators. The first is the

within-2SLS estimator ($W2SLS$), which transforms the estimation model, deviating from individual means, wiping out the individual effects c_i. The $W2SLS$ is consistent with simply exogenous instruments (the individual effect is correlated with the covariates). However, it is inefficient with doubly exogenous variables (no correlation between individual effects and covariates). Furthermore, it does not take into account the effects of time-invariant factors.

The second possibility is to consider the error component instrumental variables estimator ($EC2SLS$). This procedure was developed by Baltagi (Baltagi, 1980) (Baltagi, 1981), who applied the GLS estimator to produce consistent and efficient results with double exogenous instruments.

In both models the number of external instruments must be at least equal to the number of endogenous covariates.

We applied these techniques utilizing two different sets of instruments. In the first (IV1), the *price.tot* is regressed in the first stage with the wholesale-market price of natural gas. For the second (IV2), we used the baseline prices published by ARERA.

IV1 The external instrument used in the first instrumental model is the monthly mean of the natural gas price in the Italian wholesale market. This market is managed by GME (the Energy Services Manager), a public company owned by the Italian Ministry of the Economy and Finance.

In particular, we develop the estimation using the Intra-Day Gas Market price (*mi.gas*) as an instrument. It is calculated as the equilibrium price between natural gas supply and demand reflected in daily contracts.

The drawback of this instrument is that it does not explain distribution price elasticity. Indeed, only the selling price is correlated with the wholesale market price. Distribution is an entirely independent sector, where prices are based on the infrastructure investments carried out by Italian utilities.

$$\ln(price.tot_{i,t}) = \beta x_{i,t} + \zeta_1 \ln(mi.gas_t) + u_{i,t} \tag{3}$$

Equation 3 shows the first stage of the $2SLS$. Vector x represents the exogenous variable described in the base model (equation 2). We used the *mi.gas* price as external instrument. The latter two factors can regress respectively yearly macro trends and spatial variations that could affect the gas price.

IV2 The second instrumental model is constructed from a simulated price calculated with the baseline prices of selling, distribution, and system changes. The baseline price is what an individual pays for the first unit of natural gas.

The simulated price breaks the endogeneity for both the selling and distribution components; therefore, we expect price elasticity larger in magnitude than *Model 1* and *Model 2.IV1.*

The drawback of baseline prices is that they do not explicitly consider the block tariff system. In fact, the model does not consider the change in prices when consumers exceed a specific block.

$$\ln(price.tot_{i,t}) = \beta x_{i,t} + \xi_1 \ln(sel.base_t) + \xi_2 \ln(dis.base_t) + u_{i,t} \tag{4}$$

Equation 4 shows the first stage of the $2SLS$. The *sel.base* (the sales cost portion of the baseline price) and *dis.base* (distribution part of the baseline price) factors are the external instruments.

	Model 2.IV1				*Model 2.IV2*			
	Households		Non-households		Households		Non-households	
	W2SLS (1)	EC2SLS (2)	W2SLS (3)	EC2SLS (4)	W2SLS (5)	EC2SLS (6)	W2SLS (7)	EC2SLS (8)
ln(price.tot)	−0.2872***	−0.2339***	0.0580	0.0066	−0.5098***	−0.4769***	−0.4383***	−0.4584***
	(0.0164)	(0.0270)	(0.0731)	(0.1031)	(0.0088)	(0.0148)	(0.0507)	(0.0714)
hdd	0.4821***	0.4851***	0.4689***	0.4674***	0.4690***	0.4707***	0.4451***	0.4449***
	(0.0017)	(0.0029)	(0.0091)	(0.0131)	(0.0015)	(0.0026)	(0.0087)	(0.0125)
hum	−0.0175***	−0.0171***	−0.0044	−0.0044	−0.0181***	−0.0176***	−0.0048	−0.0046
	(0.0005)	(0.0009)	(0.0030)	(0.0043)	(0.0005)	(0.0009)	(0.0030)	(0.0043)
rain	−0.0035***	−0.0043***	−0.0137***	−0.0147**	−0.0009	−0.0015	−0.0097**	−0.0108*
	(0.0006)	(0.0010)	(0.0034)	(0.0050)	(0.0006)	(0.0010)	(0.0035)	(0.0050)
wind	−0.0116***	−0.0131***	−0.0114	−0.0149	−0.0093***	−0.0107***	−0.0068	−0.0109
	(0.0012)	(0.0020)	(0.0088)	(0.0126)	(0.0012)	(0.0020)	(0.0088)	(0.0126)
winter	0.0233***	0.0178**	0.1066***	0.1108***	0.0482***	0.0452***	0.1461***	0.1483***
	(0.0034)	(0.0059)	(0.0174)	(0.0252)	(0.0030)	(0.0052)	(0.0172)	(0.0247)
rain:winter	0.0070***	0.0071***	0.0058	0.0072	0.0065***	0.0064***	0.0054	0.0066
	(0.0008)	(0.0014)	(0.0050)	(0.0072)	(0.0008)	(0.0014)	(0.0050)	(0.0072)
winter:wind	0.0759***	0.0763***	0.0681***	0.0714***	0.0775***	0.0782***	0.0665***	0.0702**
	(0.0020)	(0.0034)	(0.0149)	(0.0216)	(0.0020)	(0.0034)	(0.0150)	(0.0216)
n.dwell		−0.1934***		−0.1746***		−0.1946***		−0.1722***
		(0.0069)		(0.0491)		(0.0070)		(0.0486)
smart.meter	−0.2202***	−0.2224***	−0.1966***	−0.1143*	−0.1991***	−0.1989***	−0.1608***	−0.0796
	(0.0048)	(0.0081)	(0.0319)	(0.0456)	(0.0046)	(0.0078)	(0.0317)	(0.0450)
self.read	−0.0532***	−0.0538***			−0.0524***	−0.0532***		
	(0.0046)	(0.0077)			(0.0047)	(0.0077)		
av.s.dwell		0.0400***		0.0114		0.0442***		0.0213
		(0.0064)		(0.0525)		(0.0065)		(0.0523)
build.age.index		−0.0824***		−0.0477		−0.0994***		−0.1327*
		(0.0088)		(0.0621)		(0.0090)		(0.0624)
welf.aid		−0.0252				−0.0248		
		(0.0259)				(0.0262)		
prod				1.3855***				1.3431***
				(0.0952)				(0.0930)
free.market	0.0201***	0.0223***	−0.1330	0.1165	0.0319***	0.0350***	−0.1283	0.1352
	(0.0038)	(0.0059)	(0.0780)	(0.1270)	(0.0037)	(0.0059)	(0.0862)	(0.1331)
gross.income	0.0380***	0.0200***	0.0229	−0.0019	0.0586***	0.0398***	0.0807***	0.0497**
	(0.0024)	(0.0036)	(0.0144)	(0.0195)	(0.0023)	(0.0035)	(0.0143)	(0.0191)
dep.index	−0.1406***	−0.1157***	−0.0892*	−0.1268*	−0.1738***	−0.1439***	−0.2377***	−0.2447***
	(0.0068)	(0.0084)	(0.0453)	(0.0549)	(0.0068)	(0.0085)	(0.0461)	(0.0558)
q1	−0.0245***	−0.0178***	0.0559***	0.0566***	−0.0482***	−0.0434***	0.0162	0.0196
	(0.0024)	(0.0040)	(0.0104)	(0.0149)	(0.0019)	(0.0033)	(0.0097)	(0.0139)
q2	−0.2692***	−0.2589***	−0.2043***	−0.2016***	−0.3046***	−0.2972***	−0.2742***	−0.2667***
	(0.0036)	(0.0060)	(0.0212)	(0.0303)	(0.0029)	(0.0050)	(0.0199)	(0.0285)
q3	−0.5060***	−0.4988***	−0.3940***	−0.3941***	−0.5328***	−0.5279***	−0.4463***	−0.4430***
	(0.0034)	(0.0057)	(0.0196)	(0.0282)	(0.0031)	(0.0052)	(0.0186)	(0.0267)
R^2	0.5567	0.5517	0.4781	0.4743	0.5544	0.5487	0.4716	0.4681
Adj. R^2	0.5443		0.4630		0.5416		0.4564	
Num. obs.	1,769,616	1,769,616	72,756	72,756	1,769,616	1,769,616	72,756	72,756
Hausman test	288.49 ($p < 0.001$)		345.53 ($p < 0.001$)		447.01 ($p < 0.001$)		342.42 ($p < 0.001$)	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 5: *Model 2* results. The natural gas price is instrumented with external instruments. Clustered standard errors in ().

For the household group, in *Model 2.IV1*, the price elasticity *price.tot* values are akin to those in *Model 1*. Indeed, the wholesale price instrument (*mi.gas*) can explain only the selling portion of the total price (*price.tot*). As in *Model 1*, the price component observed in the elasticity is the selling price. The instrument ignores distribution and system-charges, which continue to generate

endogeneity, decreasing the elasticity of the model in absolute terms.
In *Model 2.IV2*, the elasticity coefficients (*price.tot*) represent the sum of the *price.sel* and *price.dis* components. Therefore, they are larger than in *Model 1* and *Model 2.IV1*. *Model 2.IV2* results for non-households confirm the result found for *Model 1* where elasticity is weaker than in the household group. Furthermore, in *Model 2.IV1*, the price variation has no significance, indicating the negligible effect of price variations in production or commercial activities.

4.3 Discussion of Results

The study of price elasticity is the heart of this empirical study. The results show that price elasticity is negative, as expected, for all the models illustrated in the previous section. According to the literature, consumers react negatively to price variation, although the price-demand dynamic is mostly inelastic, at less than 1. This means that natural gas price changes have less impact on consumption than other factors. Table 6 summarizes price elasticities as found in the models.

	Model 1	*Model 2.IV1*	*Model 2.IV2*
Households	-0.32 ∼ -0.34	-0.23 ∼ -0.29	-0.48 ∼ -0.51
Non-households	-0.18 ∼ -0.22	-	-0.44 ∼ -0.46

Table 6: Summary of natural gas price elasticities.

The household group results indicate that a 1% increase in natural gas price reduces residential consumption in the range of $0.23 \sim 0.51\%$. This is consistent with past empirical studies, as illustrated in section 2.
The result of the non-household group is generally lower in magnitude, and the elasticity found in *Model 2.IV1* is not significant. The literature (Labandeira et al., 2012) confirms that price variations impact less on demand in production and commercial activities. However, in our sample, the non-household group is heterogeneous and includes several consumption profiles that may react to price changes.
In our opinion, natural gas use in commercial or industrial activities is marginally influenced by price changes, mostly when energy efficiency programs are ignored. However, the non-residential sample is cross-sectionally much smaller and more heterogeneous than the household group. This may lead to bias estimations caused by outlier data.
Based on the price analysis and the increase in endogeneity, we expected different elasticities from the models and demonstrate that price elasticity evolves in function of how the endogeneity is treated.
For instance, considering the residential demand, in *Model 1* and *Model 2.IV1*, endogeneity is partially considered. These models exploit only the *price.sel* component of the total price, with the lowest elasticity values in the $-0.23 \sim -0.34$ range. In *Model 2.IV2*, the aim is to fully break the endogeneity of *price.tot*. The elasticity is larger, indicating that the residential response to a price change is in the range between -0.5.

The signs of environmental variables meet our expectations. As previous empirical work has pointed out, the main variable of interest for natural gas consumption is temperature. The *hdd* is the main driver of our empirical models. The variable aims to capture the seasonality of natural gas consumption, both for residential and non-residential users. In particular, an increase in *hdd* values involves a reduction in daily mean temperature and an increase in heating. It is also worth noting that *wind* and *rain* factors are correctly positively correlated with gas demand; indeed, both cause more heat loss. However, the coefficient magnitude is dominant for the *hdd*. The other environmental factors are at least one order of magnitude lower. Especially *rain* and *hum* effects are non-significant at almost zero. For the non-household sample, the *hdd* and *wind* are the only two significant weather factors. Weather results have no significant differences between the various static methods and groups.

Home characteristics are important factors for residential natural gas consumption. In particular, in northern Italy, the most common heating source is natural gas. The size of the home, construction materials, and type of appliances are essential information significantly influencing energy demand. Unfortunately, this information is not available in our data sets. Therefore, we sought to mitigate the omitted variable effects with several dummies and socioeconomic macro trends.
The number of homes in the same building (*n.dwell*) is obtained from the address information in the distribution data set. To the best of our knowledge, this variable is tested for the first time. We started from the assumption that single houses are generally larger with more heat loss. The results confirm this hypothesis, relating natural gas demand negatively with the *n.dwell* variable in all the models. This factor is also relevant to the non-household group, probably due to the commercial activities in the sample with similar characteristics to the residential group.
The macro trends of home characteristics are meaningful, especially for households. The variation of the average size of the home (*av.s.dwell*) and the building age index (*build.age.index*) accurately show that areas with larger and newer homes consume respectively more and less natural gas.
Another significant factor in the regressions is the *smart.meter*, used to read consumption. The main reason for installing smart-metering technology with end-users is to guarantee billing based on actual consumption and not on estimates. Estimated bills must be adjusted in subsequent invoices after an actual reading, so bills do not reflect real consumption. As observed by Gans et al. (Gans et al., 2013), the provision of the additional information by smart-meters reduces energy consumption (natural gas in our case) by end-users. Consumers with accurate billing invoices can check and adjust their consumption. The negative sign for the smart-metering dummy is consistent with this assumption. Furthermore, the issue has been extensively studied in the electricity market (Schleich et al., 2012) (Carroll et al., 2014), with similar results.
Similarly, the self-reading dummy (*self.read*) should be considered, negatively correlating with natural gas demand. The self-reading of gas meters produces accurate billing invoices that do not need future adjustments by providers (except for reading errors by end-users). Furthermore, consumers who read their own meters are likely to be more aware of, and sensitive, to cost.

The socioeconomic factors considered for the residential cluster are: macro trends indicating

yearly gross income per capita and dependent population index ($gross.income$ and $dep.index$), the market-liberalization dummy ($free.market$), and the indication of low-income households ($welf.aid$). For the non-household group, the low-income indicator (relevant only for households) is replaced by the dummy variable $prod$, indicating whether natural gas is also used for manufacturing.
The macro trends introduce the spatial variability of the sample; since the data are annual, temporal variability is limited. The resulting signs are consistent with our expectations. The income indicator ($gross.income$) is positive; locations with high incomes are assumed to consume more energy. The $dep.index$ is negative, indicating that the population tends to consume more in the working-age. Finally, macro trends can be useful to mitigate the lack of individual information; the influence of these socioeconomic variables is small but significant.
The dummy ($welf.aid$) shows that the population that received energy welfare subsidies in 2016-2018 tends to consume slightly less than the rest of the sample. However, the difference is small in all the models, indicating that the effect is limited. A similar consideration applies to the liberalization process dummy ($free.market$), where the magnitude does not indicate a substantial difference between free-market and CPM users. Furthermore, the positive sign of the residential group indicates that market deregulation does not influence consumer sensitivity to energy consumption. This result is consistent, for example, with the paper of Nakajima et al. (Nakajima and Hamori, 2010) who found no substantial difference in price elasticity between deregulated and non-deregulated states in the US.

Finally, the research on price elasticity might have several practical implications. For instance, we have observed how households react to price changes in the study area. A significant increase in prices would lead households to use energy more efficiently. This can be particularly useful for local policymakers, especially in the short term when a reduction in energy consumption is necessary. In the area of study is fundamental to study proper administrative tools to restrict the level of emissions[10]. Furthermore, from our results, the level of technological innovation in the energy grid ($smart.meter$ and $self.read$ dummies) has a clear influence on residential gas consumption. This factor could be exploited by energy authorities for a long-term demand reduction, stimulating awareness of energy consumption in the population.

Our results could be important for energy utilities, as well. The use of price elasticity can be a useful tool to forecast companies' revenues or to design energy scenarios. The topic is nowadays particularly important due to the huge volatility of spot and forward natural gas and electricity prices that can significantly influence the evolution of local energy demand and, consequently, utilities' revenues. Furthermore, we have shown how companies and large consumers are hardly affected by the variations in prices, with elasticities always lower than residential end-users. This information could be useful to set proper price policies.

The study of price elasticity can also be a fundamental tool to design proper regulations and incentives. For example, (Xiang and Lawley, 2019) estimate the effects of the 2008 carbon tax in the British Columbia on residential natural gas consumption per capita. Implementing a panel data

[10] Vicenza air quality is classified on average as "poor" by European Environment Agency. https://www.eea.europa.eu/themes/air/urban-air-quality/european-city-air-quality-viewer

analysis, they use several input variables with data from 1990 to 2014, such as gas demand, weather conditions, fuel prices, building and population characteristics. Their results suggest that factors such as weather and household characteristics have a crucial impact in the short and medium-term, compared to gas and household incomes. However, they are able to assess the long-term impacts of a carbon tax. The overall results suggest that a carbon tax of 1 cent/m3 reduces per capita natural gas consumption by 3%. Hence, the result can be interpreted as negative price-demand elasticity. In a very recent paper, (Liu et al., 2022) study the effect of a carbon tax on the demand of several fuels in China. The paper explores the influence that energy price increases implied by a carbon tax has on the on fuel consumption using a quadratic almost ideal demand system (QUAIDS) model. According to the output of the estimated model, the authors state the impact of carbon tax on natural gas is neutral at aggregate national level and in urban areas, while a negative progressive elasticity is observed in rural areas. Price elasticity of energy consumption could have policy implications in the field of energy efficiency. The effect of a carbon tax on energy efficiency has been indeed studied by (Adetutu et al., 2020). The methodology applied by the authors relies on a procedure for the decomposition of the energy intensity based on a stochastic energy frontier. They collect data on a sample of firms in UK and find that the impact of a carbon tax on energy efficiency is lower than the substitution effect among fuels.

5 Conclusion

This paper shows how consumers respond to price variations when the authorities liberalize the retail energy sector. In particular, we focus on natural gas, which will be fundamental in the near future for local strategies regarding energy demand and decarbonization.

In the Italian sector, household natural gas consumption deserves the attention of policymakers, since it is the primary thermal energy source. During the period of analysis, households had few alternatives to natural gas for heating. Indeed, the electricity is generally much more expensive than natural gas and the district heating networks are limited in extension. In 2013, natural gas was the primary source for residential heating (72.3 %) in the Veneto region. Electricity was used for heating only by the 1.5 % of households [11]. However, in recent years, structural and plant redevelopment of residential buildings have been encouraged by the Italian government. This should lead to an increasing use of heating pumps for domestic hot water (33.3 % of new buildings) and heating (40 % of new buildings) (Regione Veneto, 2017).

The international literature includes numerous empirical cases investigating the factors driving energy consumption. However, natural gas distribution is studied marginally compared to electricity, especially when disaggregated data are available.

To the best of our knowledge, our paper includes the first application of panel models to Italian natural gas billing data at the individual level. Considering this kind of data, we adopt a bottom-up statistical approach analyzing each individual contribution and studying the principal economic and environmental drivers. Similar techniques have been applied in other countries (see section 2); however, our analysis distinguishes itself due to the integration of billing and distribution data

[11] ISTAT, 2013. Consumi energetici delle famiglie, `http://dati.istat.it/Index.aspx?QueryId=22809`

(which are usually private and difficult to obtain), as well as weather factors. Furthermore, we analyze the different types of endogeneity in the price-demand relationship in detail, suggesting possible statistical solutions.

Previous analyses regarding the Italian case never combined disaggregated micro-data (i.e., bililng data) with elesticitiy estimation. For instance, Besagni et al. (2018) use disaggregated data, but not billing data, and do not estimate price elasticities. On the other hand, Bianco et al. (2014) estimate price elasticities applying a top-down approach without using billing data.

Given the local nature of the data, our study is initially of local interest in the area of Vicenza in northern Italy but turns out to have widespread policy implications. The Veneto region is the core of Italian manufacturing with a primary role in national energy consumption. Moreover, our methods transcend spatial and temporal limitations and can be applied in any region or country.

To study natural gas price elasticity and consumption patterns, we assembled three-panel data models, collecting data from different sources in the period January 2016-December 2018: (1) monthly billing data, (2) weather factors, (3) technical distribution data, and (4) macro socioeconomic trends. The billing and distribution data are interpolated using an identifier code for each end-user. The weather factor and macro-trends relate to each ZIP code.

Macro-trends include average gross income (yearly variation), the average size of the home (fixed over time), the building-age index (fixed over time), and the dependence-population index (yearly variation). The combination of these factors with micro-data are useful to overcome the lack of individual information in the data-set and mitigate the endogeneity effect due to the omitted variable problem. The signs of regression coefficients are significant and coherent with our expectation (e.g., wealthier areas with larger homes consume more natural gas.); however their magnitude is generally lower than that related to other variables, such as the main drivers of gas consumption (e.g., temperature), the presence of smart meters, the number of dwellings in the same building. The panel models solves the increase in endogeneity with different methods. In *Model 1*, the price component is a demand-function is eliminated. In *Model 2*, price elasticity as studied considering instrumental variable methods. In particular, the wholesale market price and the baseline prices established quarterly by ARERA are used as external instruments to break the price endogeneity. Furthermore, in Appendix B, a dynamic panel data model is proposed with the first lag of natural gas consumption in the regressor matrix (this model is deeply analyzed in Favero's doctoral thesis(Favero, 2022)). Of the static models our preference is for fixed-effect methods because Hausman tests suggest a correlation between the individual fixed effect and the regressors, rendering RE methods inconsistent. However, since the results are similar, we also used RE estimates to interpret the factors fixed over time.

The results obtained from the models have the signs expected and are robust across the different static models. However, the dynamic panel model defined in Appendix B has some inconsistencies in correlation signs, indicating that the consumption lag in the explanatory variables significantly affects the regression coefficients.

The estimates of price elasticity appear to be consistent across the various specifications and the previous literature. The values are between -0.2 and -0.5 for households and -0.2 and -0.4 for non-households, based on the level of endogeneity resolution. However, some price components are unobserved in our investigation. Indeed, in addition to VAT, end-users have fixed costs irrespective of gas consumption. Furthermore, end-users are subject to billing adjustments after the meter reading and different invoice periods (monthly or bimonthly). Unfortunately, due to the specificity of each end-user, it was not possible to construct a complete data-set including all the possible cases.

The analyses and results illustrated in this paper should be interpreted with caution. We estimated price elasticity in a relatively small area and for one utility only, as seller and distributor. Therefore, the outcomes might be biased by the sample selection. The analysis can be considered a reference point for the population of northern Italy but is less significant if the entire Italian population is considered. Furthermore, Agsm Aim Energia Srl is the only source of billing data and the data could therefore be affected by unobserved structural collection errors.

The large data-set (especially in the household group) and cluster standard errors ensure consistency even in the presence of individual heterogeneity. Moreover, elasticity increases with the endogeneity resolution level. This can be considered proof of the accuracy of the estimates. Of the proposed models, *Model 2.IV2* is preferable, where the total price is instrumented directly from the prices established by ARERA.

The dynamic model proposed in Appendix B deserves further study; in particular, the consumption lag value seems to be a source of distortion, overestimating price elasticity and changing the sign of other factors such as the use of a smart-meter and the effects of the wind and humidity.

The lack of individual information is mitigated in our models by using macro trends and quarter dummies. Further research could integrate the billing and distribution data sets with a survey of the individual characteristics of consumers. The significance level of macro trends confirms the importance of home characteristics, income, and household composition. Furthermore, while we have conducted a simple simulation of the effects of market liberalization during the current transition phase, a more detailed analysis of liberalization will be possible when it has been completed. Similarly, the spread of smart-metering technology should be analyzed in depth to study the implications for consumers and exploit this possible new source of data.

Acknowledgments

The authors wish to thank the Associate Editor and two anonymous referees for their helpful comments and suggestions. Agsm Aim Energia and V-Reti (Agsm Aim group) are gratefully acknowledged for their assistance, expertise, and for providing the billing and distribution data-sets. However, results and interpretations illustrated in the paper do not reflect the opinion of the Agsm Aim group, but should be uniquely attributed to the authors. Filippo Favero wishes to thank the attendees of a seminar delivered on 8th September 2020 at the Department of Economics, University of Verona, where preliminary results of the manuscript were presented. Luigi Grossi thanks the attendees of the SIS conference, that has taken place in Caserta on June 2022, where extensions of the present paper have been presented and discussed. Some ideas has been used to revise the present manuscript. The research has been partially funded by the Italian Ministry of University (grant FFABR2017). The usual disclaimer applies.

References

Adetutu, M. O., Odusanya, K. A., and Weyman-Jones, T. G. (2020). Carbon tax and energy intensity: Assessing the channels of impact using UK microdata. *The Energy Journal*, 41:143–166.

Alberini, A. and Filippini, M. (2011). Response of residential electricity demand to price: The effect of measurement error. *Energy Economics*, 33(5):889–895.

Alberini, A., Gans, W., and Velez-Lopez, D. (2011). Residential consumption of gas and electricity in the U.S.: The role of prices and income. *Energy Economics*, 33(5):870–881.

Alberini, A., Khymych, O., and Scasny, M. (2020). Responsiveness to energy price changes when salience is high: Residential natural gas demand in Ukraine. *Energy Policy*, 144:111534.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2):277–297.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29 – 51.

Auffhammer, M. (2022). Climate adaptive response estimation: Short and long run impacts of climate change on residential electricity and natural gas consumption. *Journal of Environmental Economics and Management*, 114:102669.

Auffhammer, M. and Rubin, E. (2018). Natural gas price elasticities and optimal cost recovery under consumer heterogeneity: Evidence from 300 million natural gas bills. Working Paper 24295, National Bureau of Economic Research.

Baltagi, B. (2005). *Econometric analysis of panel data*. Wiley, Chichester [u.a.], 3. ed. edition.

Baltagi, B. H. (1980). On Seemingly Unrelated Regressions with Error Components. *Econometrica*, 48(6):1547–1551.

Baltagi, B. H. (1981). Simultaneous equations with error components. *Journal of Econometrics*, 17(2):189 – 200.

Batalla-Bejerano, J., Costa-Campi, M. T., and Trujillo-Baute, E. (2016). Collateral effects of liberalisation: Metering, losses, load profiles and cost settlement in Spain's electricity system. *Energy Policy*, 94(C):421–431.

Batalla-Bejerano, J., Trujillo-Baute, E., and Villa-Arrieta, M. (2020). Smart meters and consumer behaviour: Insights from the empirical literature. *Energy Policy*, 144:111610.

Besagni, G. and Borgarello, M. (2018). The determinants of residential energy expenditure in Italy. *Energy*, 165:369–386.

Bianco, V., Scarpa, F., and Tagliafico, L. A. (2014a). Analysis and future outlook of natural gas consumption in the Italian residential sector. *Energy Conversion and Management*, 87:754–764.

Bianco, V., Scarpa, F., and Tagliafico, L. A. (2014b). Scenario analysis of nonresidential natural gas consumption in Italy. *Applied Energy*, 113:392–403.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115 – 143.

Boogen, N., Datta, S., and Filippini, M. (2017). Dynamic models of residential electricity demand: Evidence from Switzerland. *Energy Strategy Reviews*, 18:85–92.

Brabec, M., Konár, O., Pelikán, E., and Malý, M. (2008). A nonlinear mixed effects model for the prediction of natural gas consumption by individual customers. *International Journal of Forecasting*, 24(4):659–678.

Breusch, T. S. and Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47(5):1287–1294.

Burke, P. J. and Yang, H. (2016). The price and income elasticities of natural gas demand: International evidence. *Energy Economics*, 59:466–474.

Carroll, J., Lyons, S., and Denny, E. (2014). Reducing household electricity demand through smart metering: The role of improved information about energy saving. *Energy Economics*, 45:234 – 243.

Costa-Campi, M. T. and Trujillo-Baute, E. (2015). Retail price effects of feed-in tariff regulation. *Energy Economics*, 51:157–165.

Croissant, Y. and Millo, G. (2018). *Panel Data Econometrics with R: the plm package*. Wiley.

Dong, K., Dong, X., and Sun, R. (2019). How did the price and income elasticities of natural gas demand in China evolve from 1999 to 2015? The role of natural gas price reform. *Petroleum Science*, 16:685–700.

Favero, F. (2022). Empirical investigation of natural gas consumption and price elasticity: evidence from billing and distribution data in Italy. *Doctoral Thesis University of Verona*.

Fell, H., Li, S., and Paul, A. (2014). A new look at residential electricity demand using household expenditure data. *International Journal of Industrial Organization*, 33(1):37–47.

Filippini, M. and Kumar, N. (2021). Gas demand in the Swiss household sector. *Applied Economics Letters*, 28(5):359–364.

Fumo, N. and Rafe Biswas, M. A. (2015). Regression analysis for prediction of residential energy consumption. *Renewable and Sustainable Energy Reviews*, 47:332–343.

Gans, W., Alberini, A., and Longo, A. (2013). Smart meter devices and the effect of feedback on residential electricity consumption: Evidence from a natural experiment in Northern Ireland. *Energy Economics*, 36:729–743.

Gürsan, C. and de Gooyert, V. (2021). The systemic impact of a transition fuel: Does natural gas help or hinder the energy transition? *Renewable and Sustainable Energy Reviews*, 138:110552.

Halvorsen, B. and Larsen, B. M. (2013). How serious is the aggregation problem? An empirical illustration. *Applied Economics*, 45(26):3786–3794.

Kostakis, I., Lolos, S., and Sardianou, E. (2021). Residential natural gas demand: Assessing the evidence from Greece using pseudo-panels, 2012–2019. *Energy Economics*, 99:105301.

Labandeira, X., Labeaga, J. M., and López-Otero, X. (2012). Estimation of elasticity price of electricity with incomplete information. *Energy Economics*, 34(3):627–633.

Labandeira, X., Labeaga, J. M., and López-Otero, X. (2017). A meta-analysis on the price elasticity of energy demand. *Energy Policy*, 102:549–568.

Liddle, B. and Huntington, H. (2021). How prices, income, and weather shape household electricity demand in high-income and middle-income countries. *Energy Economics*, 95:104995.

Liu, G., Dong, X., Jiang, Q., Dong, C., and Li, J. (2018). Natural gas consumption of urban households in china and corresponding influencing factors. *Energy Policy*, 122:17 – 26.

Liu, J., Gong, N., and Qin, J. (2022). How would the carbon tax on energy commodities affect consumer welfare? evidence from China's household energy consumption system. *Journal of Environmental Management*, 317:115466.

McIntyre, S. (2018). Exploring households' responsiveness to energy price changes using microdata. Working Papers 1806, University of Strathclyde Business School, Department of Economics.

Meier, H. and Rehdanz, K. (2010). Determinants of residential space heating expenditures in Great Britain. *Energy Economics*, 32(5):949–959.

Miller, M. and Alberini, A. (2016). Sensitivity of price elasticity of demand to aggregation, unobserved heterogeneity, price trends, and price endogeneity: Evidence from U.S. Data. *Energy Policy*, 97:235–249.

Nakajima, T. and Hamori, S. (2010). Change in consumer sensitivity to electricity prices in response to retail deregulation: A panel empirical analysis of the residential demand for electricity in the United States. *Energy Policy*, 38(5):2470–2476.

Okajima, S. and Okajima, H. (2013). Estimation of Japanese price elasticities of residential electricity demand, 1990–2007. *Energy Economics*, 40:433 – 440.

Opolska, I. (2017). The efficacy of liberalization and privatization in introducing competition into European natural gas markets. *Utilities Policy*, 48:12–21.

Otero, J. and Waddams Price, C. (2001). Price discrimination, regulation and entry in the UK residential electricity market. *Bulletin of Economic Research*, 53(3):161–175.

Pereira Uhr, D. d. A., Squarize Chagas, A. L., and Ziero Uhr, J. G. (2019). Estimation of elasticities for electricity demand in Brazilian households and policy implications. *Energy Policy*, 129:69–79.

Ravnik, J. and Hriberšek, M. (2019). A method for natural gas forecasting and preliminary allocation based on unique standard natural gas consumption profiles. *Energy*, 180:149–162.

Regione Veneto (2017). Piano Energetico Regionale, Fonti Rinnovabili Risparimio Energetico Efficienza Energetica.

Safari, A., Das, N., Langhelle, O., Roy, J., and Assadi, M. (2019). Natural gas: A transition fuel for sustainable energy system transformation? *Energy Science & Engineering*, 7(4):1075–1094.

Schleich, J., Klobasa, M., and Gölz, S. (2012). Does smart metering reduce residential electricity demand? In *2012 9th International Conference on the European Energy Market*, pages 1–4.

Silva, S., Soares, I., and Pinho, C. (2017). Electricity demand response to price changes: The Portuguese case taking into account income differences. *Energy Economics*, 65:335 – 342.

Soldo, B. (2012). Forecasting natural gas consumption. *Applied Energy*, 92:26–37.

Sun, C. and Ouyang, X. (2016). Price and expenditure elasticities of residential energy demand during urbanization: An empirical analysis based on the household-level survey data in China. *Energy Policy*, 88:56–63.

Swan, L. G. and Ugursal, V. I. (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, 13(8):1819–1835.

Trujillo-Baute, E., del Río, P., and Mir-Artigues, P. (2018). Analysing the impact of renewable energy regulation on retail electricity prices. *Energy Policy*, 114:153–164.

Valor, E., Meneu, V., and Caselles, V. (2001). Daily air temperature and electricity load in Spain. *Journal of Applied Meteorology*, 40(8):1413 – 1421.

Woo, C. K., Liu, Y., Zarnikau, J., Shiu, A., Luo, X., and Kahrl, F. (2018). Price elasticities of retail energy demands in the United States: New evidence from a panel of monthly data for 2001–2016. *Applied Energy*, 222:460–474.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data.* MIT Press, Cambridge and London.

Xiang, D. and Lawley, C. (2019). The impact of British Columbia's carbon tax on residential natural gas consumption. *Energy Economics*, 80:206–218.

Zeng, S., Chen, Z. M., Alsaedi, A., and Hayat, T. (2018). Price elasticity, block tariffs, and equity of natural gas demand in China: Investigation based on household-level survey data. *Journal of Cleaner Production*, 179:441–449.

Zhang, Y., Ji, Q., and Fan, Y. (2018). The price and income elasticity of China's natural gas demand: A multi-sectoral perspective. *Energy Policy*, 113:332–341.

Zhu, X., Li, L., Zhou, K., Zhang, X., and Yang, S. (2018). A meta-analysis on the price elasticity and income elasticity of residential electricity demand. *Journal of Cleaner Production*, 201:169–177.

Appendix A Robustness Check

To test the robustness of the panel analysis, we carried out a robustness check excluding possible outliers from the estimation. Therefore, we filtered the PODs with natural gas consumption (Q) lower than the 1st percentile and higher than the 99th percentile. Then, we estimated the *Model 2* for the new data-set, comparing the results in tables A1 and 5. The difference in price elasticity is marked for Households in $Model2.IV2$, however, other coefficients and other models do not present relevant differences, suggesting a negligible effects of outliers.

	Model 2.IV1				*Model 2.IV2*			
	Households		Non-households		Households		Non-households	
	W2SLS (1)	EC2SLS (2)	W2SLS (3)	EC2SLS (4)	W2SLS (5)	EC2SLS (6)	W2SLS (7)	EC2SLS (8)
ln.price.tot	-0.2699^{***}	-0.2114^{***}	0.1068	0.0708	-0.6197^{***}	-0.5767^{***}	-0.4468^{***}	-0.4493^{***}
	(0.0164)	(0.0280)	(0.0685)	(0.1047)	(0.0095)	(0.0164)	(0.0508)	(0.0772)
hdd	0.4436^{***}	0.4447^{***}	0.4526^{***}	0.4515^{***}	0.4349^{***}	0.4356^{***}	0.4406^{***}	0.4401^{***}
	(0.0013)	(0.0024)	(0.0079)	(0.0124)	(0.0013)	(0.0023)	(0.0078)	(0.0121)
hum	-0.0129^{***}	-0.0128^{***}	-0.0008	-0.0002	-0.0106^{***}	-0.0102^{***}	0.0036	0.0042
	(0.0005)	(0.0009)	(0.0028)	(0.0044)	(0.0005)	(0.0009)	(0.0028)	(0.0043)
rain	-0.0095^{***}	-0.0099^{***}	-0.0217^{***}	-0.0221^{***}	-0.0069^{***}	-0.0071^{***}	-0.0182^{***}	-0.0186^{**}
	(0.0007)	(0.0013)	(0.0042)	(0.0066)	(0.0008)	(0.0013)	(0.0042)	(0.0066)
winter	0.0974^{***}	0.0947^{***}	0.0185	0.0232	0.1147^{***}	0.1128^{***}	0.0558^{**}	0.0576
	(0.0036)	(0.0064)	(0.0203)	(0.0318)	(0.0036)	(0.0063)	(0.0197)	(0.0308)
wind	0.0082^{***}	0.0074^{*}	-0.0051	-0.0081	0.0068^{***}	0.0056	-0.0064	-0.0095
	(0.0017)	(0.0030)	(0.0113)	(0.0177)	(0.0017)	(0.0030)	(0.0114)	(0.0177)
rain:winter	0.0271^{***}	0.0270^{***}	0.0198^{***}	0.0199^{*}	0.0262^{***}	0.0258^{***}	0.0171^{**}	0.0171^{*}
	(0.0009)	(0.0016)	(0.0052)	(0.0082)	(0.0009)	(0.0016)	(0.0053)	(0.0082)
winter:wind	0.0102^{***}	0.0092^{*}	0.0302	0.0328	0.0206^{***}	0.0204^{***}	0.0391^{*}	0.0413
	(0.0026)	(0.0046)	(0.0162)	(0.0254)	(0.0026)	(0.0046)	(0.0164)	(0.0255)
n.dwell		-0.1371^{***}		-0.1840^{***}		-0.1381^{***}		-0.1817^{***}
		(0.0064)		(0.0489)		(0.0066)		(0.0485)
smart.meter	-0.1955^{***}	-0.1983^{***}	-0.1535^{***}	-0.0829	-0.1614^{***}	-0.1618^{***}	-0.1131^{***}	-0.0447
	(0.0048)	(0.0083)	(0.0290)	(0.0453)	(0.0046)	(0.0081)	(0.0290)	(0.0448)
self.read	-0.0568^{***}	-0.0580^{***}			-0.0557^{***}	-0.0572^{***}		
	(0.0049)	(0.0084)			(0.0049)	(0.0084)		
av.s.dwell		0.0368^{***}		0.0096		0.0416^{***}		0.0165
		(0.0063)		(0.0515)		(0.0064)		(0.0516)
build.age.index		-0.0927^{***}		-0.0297		-0.1205^{***}		-0.1306^{*}
		(0.0089)		(0.0637)		(0.0092)		(0.0649)
welf.aid		-0.0040				-0.0040		
		(0.0248)				(0.0253)		
prod				1.0906^{***}				1.0572^{***}
				(0.0958)				(0.0938)
free.market	0.0219^{***}	0.0230^{***}	-0.1269	0.0979	0.0413^{***}	0.0431^{***}	-0.1252	0.1129
	(0.0038)	(0.0062)	(0.0919)	(0.1355)	(0.0038)	(0.0061)	(0.1036)	(0.1436)
gross.income	0.0394^{***}	0.0230^{***}	0.0209	-0.0029	0.0692^{***}	0.0502^{***}	0.0790^{***}	0.0501^{*}
	(0.0024)	(0.0037)	(0.0140)	(0.0204)	(0.0024)	(0.0037)	(0.0141)	(0.0204)
dep.index	-0.1585^{***}	-0.1180^{***}	-0.0879^{*}	-0.1145^{*}	-0.2156^{***}	-0.1617^{***}	-0.2676^{***}	-0.2537^{***}
	(0.0070)	(0.0085)	(0.0432)	(0.0559)	(0.0072)	(0.0088)	(0.0448)	(0.0578)
q1	0.0140^{***}	0.0182^{***}	0.1239^{***}	0.1273^{***}	-0.0028	0.0011	0.1073^{***}	0.1121^{***}
	(0.0025)	(0.0045)	(0.0128)	(0.0202)	(0.0025)	(0.0045)	(0.0130)	(0.0203)
q2	-0.2294^{***}	-0.2221^{***}	-0.1962^{***}	-0.1908^{***}	-0.2653^{***}	-0.2592^{***}	-0.2309^{***}	-0.2241^{***}
	(0.0036)	(0.0064)	(0.0177)	(0.0280)	(0.0035)	(0.0062)	(0.0181)	(0.0283)
q3	-0.4370^{***}	-0.4330^{***}	-0.3504^{***}	-0.3469^{***}	-0.4590^{***}	-0.4559^{***}	-0.3639^{***}	-0.3604^{***}
	(0.0046)	(0.0082)	(0.0242)	(0.0382)	(0.0047)	(0.0083)	(0.0244)	(0.0382)
R^2	0.5624	0.5569	0.4781	0.4857	0.5576	0.5520	0.4777	0.4739
Adj. R^2	0.5499		0.4630		0.5450		0.4627	
Num. obs.	1,540,044	1,540,044	64,692	64,692	1,540,044	1,540,044	64,692	64,692

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table A1: *Model 2* results. The natural gas price is instrumented with external instruments. Clustered standard errors in ().

Appendix B Dynamic Panel Model

The models implemented in the main part of the paper are static; the correlation with past consumption has never been considered. A possible development is to implement a dynamic panel model characterized by delayed dependent variables within the regressor matrix (equation 5).

$$\ln(Q_{i,t}) = \theta \ln(Q_{i,t-1}) + \alpha \ln(p_{i,t}) + \beta w_{r,t} + \gamma d_{i,t} + \phi e_{r,t} + \lambda q_t + c_i + u_{i,t} \tag{5}$$

where θ is the parameter relating to the autoregressive component.

Implementing models with both individual effects and lagged dependent variables can significantly reduce omitted variable endogeneity.

As estimator we used the Blundell-Bond estimator (sys-GMM) (Blundell and Bond, 1998) (Arellano and Bover, 1995), an expansion of the classic Arellano-Bond estimator (GMM) (Arellano and Bond, 1991). The main reason for preferring the Blundell-Bond estimator is to limit weak instruments that can affect the classic Arellano-Bond approach. Furthermore, in our model, the number of times is sufficient (more than 10) to limit the number of instrument lags to 10, avoiding the propagation of weak instruments.

By construction, in GMM and sys-GMM models, residuals are affected by serial correlation of the first order; however, residuals should not exhibit a significant serial correlation of the second order.

Table B1 shows the results of the dynamic panel models. Unlike the previous models, the variables considered in the dynamic panel model exclude the macro trends and the constant dummies over time that cause singularities.

The regression coefficients confirm that residential is more responsive than non-residential consumption. Furthermore, the elasticities found with the sys-GMM estimator have a greater magnitude than the previous model. Unlike *Model 1* and *Model 2*, the use of a dynamic model can potentially entirely solve the endogeneity problem for both the price components (*price.sel* and *price.dis*) and for the block tariff system. Indeed, exploiting past consumption as an instrument can define individual consumption patterns, identifying past correlations between consumption and variations in the price as a demand function (the second type of endogeneity). However, other coefficients (e.g., *hdd*, *wind*, *smart.meter*) are different in magnitude and sign than the static models. It is reasonable to assume that the lag of gas consumption captures much of the variability, yielding less significant and possible inconsistent outcomes.

	Model 3	
	Households	Non-households
	sys-GMM (1)	sys-GMM (2)
lag(ln(Q), 1)	0.9226***	0.9287***
	(0.0017)	(0.0104)
ln(price.tot)	−0.7404***	−0.4695***
	(0.0075)	(0.0468)
hdd	−0.0030***	0.0431***
	(0.0016)	(0.0088)
wind	−0.0450***	−0.0445**
	(0.0010)	(0.0058)
rain	0.00220***	0.0216***
	(0.0005)	(0.0027)
hum	−0.0077**	0.0075***
	(0.0008)	(0.0042)
smart.meter	0.2111***	0.1243***
	(0.0043)	(0.0224)
self.read	−0.0072*	
	(0.0038)	
free.market	0.0299***	0.2030***
	(0.0019)	(0.0285)
$q1$	−0.4547***	−0.3666***
	(0.0029)	(0.0180)
$q2$	−0.8329***	−0.6364***
	(0.0041)	(0.0247)
$q3$	−0.3868***	−0.2370***
	(0.0046)	(0.0233)
Autocorrelation test (2nd order)	0.8037 ($p = 0.9359$)	0.4013 ($p = 0.6882$)
Num. obs.	1,769,616	72,756

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table B1: Dynamic Panel model with Blundell Bond estimator. Clustered standard errors in ().

The authors wish to acknowledge the time devoted by the Editor, Associate Editor and Reviewer#1 to the second review of the paper. Many thanks to Reviewer#1 for the new suggestions. The authors have taken them into account in the new revision of the manuscript. They think that the new version has improved the previous one and hope that the current version of the manuscript could be positively evaluated.

In this document comments made by Reviewer#1 are in italic font and the responses in normal font.

Reviewer #1: I am glad that many of the reviewers' comments have been addressed. However, I think the paper needs some more (minor) revisions before proceeding to a possible publication. In particular:

Point 1) Policies should be very specific and clear.

Section 4 has been expanded to include new more specific policy implications of our results. In particular, we have commented the results of the empirical analysis and suggested possible policy implications for the management of gas supply at the local level (see lines 605-623 of the new version).

Point 2) Regarding the updated Table 1, authors could take into account some of the following publications:

- Filippini, M., & Kumar, N. (2021). Gas demand in the Swiss household sector. Applied Economics Letters, 28(5), 359-364.

- Kostakis, I., Lolos, S., & Sardianou, E. (2021). Residential natural gas demand: Assessing the evidence from

Greece using pseudo-panels, 2012-2019. Energy Economics, 99, 105301.

- Alberini, A., Khymych, O., & Ščasný, M. (2020). Responsiveness to energy price changes when salience is high: Residential natural gas demand in Ukraine. Energy Policy, 144, 111534.

- Dong, K., Dong, X., & Sun, R. (2019). How did the price and income elasticities of natural gas demand in China evolve from 1999 to 2015? The role of natural gas price reform. Petroleum Science, 16(3), 685-700.

- Zhang, Y., Ji, Q., & Fan, Y. (2018). The price and income elasticity of China's natural gas demand: A multisectoral perspective. Energy Policy, 113, 332-341.

Many thanks for these new references. We have included them in the reference list, with exception of "Alberini, A., Khymych, O., and Scasny, M. (2020)", which was already included in the previous version. We have updated Table 1 accordingly summarizing the main results contained in each paper.

ELSEVIER ECONOMICS INCLUSION AND DIVERSITY FORM

Inclusion and diversity are at the heart of the way we work, think, and run our business at Elsevier. We are committed to making a difference in the realm of academia and publishing.

With this in mind, we would like to assess inclusion and diversity initiatives and their impact, and need information about you, our authors, and about your practices. We would like to therefore ask you to complete this Inclusion and Diversity form. **The information you provide in this form will not have any influence on the scientific consideration of your paper.**

The purpose of this inclusion and diversity initiative is similar to existing practices in publishing, such as Declaration of Interests, Author Contributions, and Data and Code Availability. The form encompasses a wide variety of aspects, including research methodology, sample selection, and (co-)author profiles. Based upon the answers you provide, you will have the option to include an Inclusion and Diversity statement in your paper. By adding a statement to your paper you would have the opportunity to increase transparency, raise awareness concerning inclusion and diversity in the world of academia, and highlight your publication as a best practice example. **If you do not wish to participate or publish such a statement alongside your paper, please indicate this preference at the top of page 2 of this form.**

<u>What types of articles does this apply to?</u>

We require one copy of this form for all original research content at the point of acceptance: research article, reports, short articles, case studies etc. You need to complete and return this form even if you choose not to provide any information.

We do not require this form for review formats (review, perspective, preview, etc).

<u>Who should complete the form?</u>

The corresponding author (or one of the corresponding authors if there is more than one) is solely responsible for completing this form, on behalf of all contributing authors.

We suggest that the corresponding author uses the opportunity of the normal process of sharing the final version of the paper with all authors before resubmission to also share this form to ask all co-authors what, if any, information they would like to share. No author is required to provide any information that they do not feel comfortable disclosing.

<u>How will the information you provide be used?</u>

We will keep the information that we gather in these forms private and secure. We will not share it with any external parties. We hope to be able to use the data for aggregated reporting at a general level about the types of inclusion and diversity efforts that our authors are making.

<u>How to return the form?</u>

Please complete this form and upload with your final submission on Editorial Manager, alongside your Declaration of Interests form (if appropriate).

If you have specific questions about how to complete this form or about this initiative, please contact your Journal Manager.

PAGE 2: INCLUSION AND DIVERSITY INFORMATION

To help us gather information in a straightforward way, this form is based on a series of options that you can choose to indicate. We understand that these are not comprehensive and apologize in advance if an effort that you would like to highlight is not included as an option. Please note that we are running this program as a pilot during the course of 2021 and that we welcome any comments and suggestions you have about this initiative and how to develop it going forward to help achieve the intended goals of promoting, fostering, and highlighting positive change in science (see page 4).

If you do not wish to provide information about inclusion and diversity, please check this box: ☐

- ***if you checked this box, please proceed to page 4 of the form***

If you do wish to participate in this initiative, please complete each section below. We recognize not all points will be relevant for every paper. Please check any/all sentences that apply to your paper:

Inclusion and diversity relating to the scientific content of the paper:

For studies involving human subjects, whether recruited (e.g. clinical analyses) or enrolled spontaneously (e.g. online surveys):

☐ We worked to ensure gender balance in the recruitment of human subjects.

☐ We worked to ensure ethnic or other types of diversity in the recruitment of human subjects.

☐ We worked to ensure that the study questionnaires were prepared in an inclusive way.

Inclusion and diversity relating to authorship and attribution:

☐ One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science.

☐ One or more of the authors of this paper self-identifies as living with a disability.

☐ One or more of the authors of this paper received support from a program designed to increase minority representation in science.

☐ While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list.

☐ The author list of this paper includes contributors from the location where the research was conducted who participated in the data collection, design, analysis, and/or interpretation of the work.*

*This check box relates to active avoidance of "helicopter science", in which authors, generally from a high-income country or non-indigenous group, rely on people and resources from a lower-income or indigenous group but then analyze and publish the data without appropriate involvement or recognition.

PAGE 3: OPTIONAL INCLUSION AND DIVERSITY STATEMENT FOR YOUR PAPER

If any of the statements on page 2 of this form apply to your paper, you have the opportunity to use them in the paper itself as an 'Inclusion and Diversity statement'.

This is an optional section that we encourage you to include as an opportunity to highlight ways in which you have taken action to promote diversity and inclusion in the context of your paper. **Your decision about whether or not to include this statement will not have any influence on the editorial handling of your paper**.

If you wish to publish an 'Inclusion and Diversity statement', please check this box: ☐

Please also verify that all authors agree to inclusion of this statement: ☐

To add an Inclusion and Diversity statement, please copy the checked sentences from page 2 that apply to your paper and include them in an 'Inclusion and Diversity statement' section, positioned in the manuscript immediately after the 'Declaration of Interests' section (please see an example below).

If you do choose to add an Inclusion and Diversity statement to your paper, please note the following:

- You will need to add this section to the paper during your final revisions, positioned after the 'Declaration of Interests' section. We will not add it for you if you do not include it yourself.
- Please only use the checked statements from page 2 exactly as written. We have purposely designed them to be phrased in a general format that will convey information in a consistent, clear and neutral way.
- If you include any information that is different from these checked statements, your Journal Manager will contact you to adjust your statement before proceeding with accepting your paper.

Inclusion and Diversity statement example:

Inclusion and Diversity

We worked to ensure sex balance in the selection of non-human subjects. One or more of the authors of this paper self-identifies as living with a disability. One or more of the authors of this paper received support from a program designed to increase minority representation in science.

PAGE 4: SIGNATURE AND COMMENTS

REQUIRED:

I declare that I have completed this form on behalf of all authors and all authors agree to the provision of any information I have indicated.

Name: Luigi Grossi

Signature:

Journal: Energy Economics

Manuscript number: ENEECO-D-22-00085

OPTIONAL:

If you wish to share any other comments or feedback, we would welcome hearing from you. For example, we would be interested to hear about any other action you may have taken to promote inclusion and diversity in the context of your paper, or suggestions you have for options that we could include in the available list of options on page 2.

Also, if you chose not to participate in this initiative, we would be interested to understand your reasoning. The information you provide in this section will not be included in the paper or our broad-based analysis but will help us in further development of this initiative and in our continuing efforts to promote Inclusion and diversity.

If you have specific questions about how to complete this form needing immediate attention, please contact your Journal Manager.

The paper relies on administrative data collected by a private commercial institution as a by-product of its main activity. Unforntunately, the authors don't have any control on the way the data are collected. For this reason, we couldn't fill out the survey in the part related to the data. However, we appreciate this initiative and support any efforts made by the scientific community in this direction.

We acknowledge the need to guarantee the right balance in the citation of genders, although sometimes it is difficult to correctly identify gender of the authors of the papers.