Applied Energy 86 (2009) 460–465

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Estimating the residential demand function for natural gas in Seoul with correction for sample selection bias

Seung-Hoon Yoo [a,*], Hea-Jin Lim [b,1], Seung-Jun Kwak [c,2]

[a] School of Business and Economics, Hoseo University, 268 Anseo-Dong, Cheonan, Chungnam 330-713, Republic of Korea
[b] Economic Research Institute, Korea Credit Guarantee Fund, 254-5 Kongdok-Dong, Mapo-Gu, Seoul 121-744, Republic of Korea
[c] Department of Economics, Korea University, 5-1 Anam-Dong, Sungbuk-Ku, Seoul 136-701, Republic of Korea

ARTICLE INFO

Article history:
Received 5 September 2007
Received in revised form 4 August 2008
Accepted 15 August 2008
Available online 18 November 2008

Keywords:
Natural gas demand
Non-response
Sample selection model

ABSTRACT

Over the last twenty years, the consumption of natural gas in Korea has increased dramatically. This increase has mainly resulted from the rise of consumption in the residential sector. The main objective of the study is to estimate households' demand function for natural gas by applying a sample selection model using data from a survey of households in Seoul. The results show that there exists a selection bias in the sample and that failure to correct for sample selection bias distorts the mean estimate, of the demand for natural gas, downward by 48.1%. In addition, according to the estimation results, the size of the house, the dummy variable for dwelling in an apartment, the dummy variable for having a bed in an inner room, and the household's income all have positive relationships with the demand for natural gas. On the other hand, the size of the family and the price of gas negatively contribute to the demand for natural gas.

1. Introduction

Over the last two decades, the consumption of natural gas in Korea has increased at a high annual growth rate of 18.7%. This is due to the expansion of the infrastructure for gas distribution and an increase in individual income. Crucial factors that underlie the recent rise in the consumption of natural gas have been an upsurge in the urban consumption of gas[3] and an increase of 55% in the proportion of the urban consumption of gas that occurs in the residential sector [1].

Owing to additional factors, in the capital city of Seoul, which accounts for 23.3% of the residential consumption of energy in Korea, a significant number of households have switched from oil and coal to gas and electricity. Especially in urban areas, where, in 1987, natural gas began to be supplied and the rate of consumption increased dramatically thereafter, the rate of consumption has increased at an average rate of 6.9% p.a. As of 2001, natural gas was the main source of energy for households in Seoul, accounting for 55.4% of the residential consumption of energy [2].

However, Korea has scarce reserves of natural gas, which is therefore imported entirely from natural gas-producing countries. The supply chain for gas in urban areas in Korea is structured as follows. Natural gas is imported in liquid form (LNG) from foreign countries and is then supplied to domestic gas companies (wholesale companies and a monopolized public utilities company that is owned by the Korean government). It is also supplied to regular, urban, gas companies (for-profit retail companies and private firms) that supply to households, businesses, and industries.

Thus, the consumer's price for urban gas is comprised of the cost of importation and the wholesale and retail firms' expenses for supplying gas. Among the various components of cost, the cost of importation fluctuates due to fluctuations in the international price of oil and the exchange rate. Being public fees, the expenses of wholesale and retail firms are regulated by local governments and legislative assemblies. In the case of Seoul, the local government suppressed price increases by lowering and freezing fees; however, after 2006, due to inflation and other factors, the price of the gas was increased for the first time [2].

On the other hand, the international price of LNG has recently been at a very high level due to the high price of oil. The high price of oil aside, the actual prices of goods are higher than ever before due to the impact of technological and seasonal factors upon the short-term supply and demand for natural gas. In particular, the price levels, which in the past several years showed a downward

* Corresponding author. Tel.: +82 41 560 8344; fax: +82 41 563 0135.
E-mail addresses: shyoo@hoseo.edu (S.-H. Yoo), iheajin@kodit.co.kr (H.-J. Lim), sjkwak@korea.ac.kr (S.-J. Kwak).
[1] Tel.: +82 2 2077 6727; fax: +82 2 2077 6722.
[2] Tel.: +82 2 3290 2217; fax: +82 2 3290 2535.

[3] The end-users of natural gas consist of the urban sector and power generation sectors. These sectors account for 63% and 37% of national consumption of gas, respectively.

0306-2619/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apenergy.2008.08.023

trend, are now showing an upward trend, as per a recently concluded long-term contract. This reverse trend is due to an increase in demand, increased uncertainty in the ability to increase supply, and the expectation that the price of oil will remain high, besides various other interdependent factors [4].

Therefore, the international LNG market is expected to be unfavorable to importing countries such as Korea that procure most of their natural gas in the form of LNG. In this context, to balance supply with demand, the demand for natural gas (NGD) needs to be accurately predicted, especially for the residential sector. This can be achieved by analyzing the pattern of residential NGD. In addition, the estimation of the demand function for natural gas will provide useful information for pricing-related policies relating to importation and the costs of supply.

Many studies have been conducted on the estimation of the residential demand for natural gas elsewhere in the world [5–15]. Most of them have been based upon the analysis of panel data but some studies [5,6,16] have used cross-sectional data. In Korea, to the best of the authors' knowledge, neither panel data sets nor investigations of the demand for natural gas are available, because of the short history of the use of LNG. Therefore, the goal of this paper is to estimate the Korean residential NGD function using cross-sectional data. Microeconomic approaches to the modeling of the residential demand for natural gas enable us to acquire a better understanding of the nature of consumer responses to a change in the price of natural gas.

To gather data, a survey was conducted of households in Seoul. As with many other cross-sectional data sets and associated surveys, the survey of households manifested a high rate of non-response [17][4]: of the 380 households sampled, 216 households (56.8%) failed to answer the question in the survey on natural gas consumption. In the present study, failure to incorporate all the missing observations was an issue.

To determine whether non-response in the present study results in bias, two questions need to be asked: Do response rates differ across identifiable groups of households? Are there systematic differences within a particular group between those who answered the questions on NGD and those who did not? Bias will occur to the extent that these between- and within-group differences exist and furthermore, are related to NGD [18].

A given cross-sectional study may suffer from either a between-group sample non-response bias and/or a within-group sample selection bias.[5] Non-response biases can generally be compensated for by weighting and imputation procedures, which assume that those participants in a given subgroup of the population who do not respond have the same expected values for NGD as those in the same subgroup who do respond. The correction procedures for those biases rely on the assumption of random non-response within the categories. In the case of sample selection bias, however, non-response is nonrandom: the individuals who do not respond to either the survey or the question hold different expected values for NGD than comparable individuals who do respond. To deal with sample selection bias, a sample selection model proposed by Heckman [19] has been commonly employed to rectify the problems emanating from the bias. However, empirical applications of the model in the estimation of the residential NGD function are lacking.

This paper, therefore, has two major purposes. The main purpose of this paper is to estimate the residential NGD function. The second purpose is to explore, during the estimation of the residential NGD function, the sample selection model that produces consistent parameter estimates and unbiased mean estimates of NGD. Moreover, the paper compares the results with those from a model that assumes no sample selection bias. It also employs two procedures, described in Section 4, to test for sample selection bias.

The remainder of the paper is organized as follows. Section 2 reviews the literature on analyses of residential NGD. Section 3 explains the data that are employed and presents an overview of the proposed methodology. The empirical findings are reported in the penultimate section, Section 4. Some concluding remarks are made in Section 5.

2. Literature review of analyses of the demand for natural gas

The empirical estimation of the NGD function has received considerable attention in the literature. Table 1 summarizes the findings. Interestingly, we could not find case-studies of Asian and African regions. Most studies focus on North American and European regions that have a long history of the consumption of natural gas. The main goal of most of these studies, especially those for North American provinces and states, was to estimate the short-run and long-run elasticities of residential NGD, for which these studies employed pooled regression.

The empirical studies can be classified into two categories, based on the data employed in them. The first category of studies uses cross-sectional survey data to analyze residential NGD. In a classic study, Verhulst [16] analyzed data from the French gas industry under the assumption that a consumer's demand for natural gas was a Cobb–Douglas function. Verhulst [16] measured the elasticities of price and disposable income. Other studies in this category include Tinic et al. [5] and Lee and Singh [6]. The latter, in particular, obtained a semi-parametric estimate considering selectivity, heteroscedasticity, and the structural stability of the demand equations for natural gas and electricity.

The second category of studies pools the data and uses panel data estimators. As noted earlier, numerous researchers have been inclined to estimate dynamic models for NGD using a time-series of cross-sections. Classic studies in this category include Balestra and Nerlove [7], who used observations on 36 US states from 1950 to 1962, and Berndt and Watkins [8], who generalized the approach of the former to a non-linear system using annual data for British Columbia and Ontario during 1959–1974. Other studies in this category include Al-Sahlawi [9], Bohi [10], Bloch [11], Beierlein et al. [12], Blatterberger et al. [13], and Herbert [14]. In a more recent study of this issue, Maddala et al. [15] applied a dynamic linear regression model using pooled data sets across 49 US states over the period of 1970–90.

In the present analysis, the estimates of the price- and income-elasticities of NGD, presented in Table 1, are of great interest. Some interesting findings emerge from the table. First, short-run price elasticity estimates of residential NGD range from −0.9 to −0.03, and long-run values range from −0.2 down to as low (or as high, in absolute value) as −4.6. According to the results of Verhulst [16], Tinic et al. [5], Al-Sahlawi [9], Bohi [10], Beierlein et al. [12], and Maddala et al. [15], both short-run and long-run price-elasticities are negative. Further, while residential NGDs are price-inelastic in the short-run, they may be price elastic (with elasticity being much larger than unity in absolute value) in the long-run.

The second interesting finding is that the estimated income-elasticities were mostly positive. These findings support the notion that natural gas is a normal good. The estimates of the short-run income-elasticities ranged from 0.01 to 0.34, and the long-run values ranged from −0.43 to 0.77. This means that NGD is

[4] Non-response can mean either refusing to fill-out and return surveys (unit non-response) or skipping items on a survey (item non-response).

[5] The term 'non-response bias,' as used in the literature on survey research, often refers to between- and within-group biases [18].