Discussion Paper

Central Bureau of Statistics, P.B. 8131 Dep, 0033 Oslo 1, Norway

No. 23 September 16, 1987

FUEL CHOICE AND THE DEMAND FOR NATURAL GAS IN WESTERN EUROPEAN HOUSEHOLDS

by

Sarita Bartlett John K. Dagsvik Øystein Olsen Steinar Strøm*

Abstract

The paper presents an econometric model for analysing natural gas demand for space heating in the residential sector in Western Europe. A discrete-continuous dynamic choice model is specified. Households' decisions on energy consumption are viewed as carried out in two steps: In the first step they choose between a limited number of fuel systems. Given this choice a decision is made on how intensively the energy equipment should be utilised.

The model is estimated from data on energy use and dwelling stock in seven European countries, organised in a database at Lawrence Berkeley Laboratory, Berkeley, USA. The results yield rather reasonable estimates on price and income elasticities of gas demand, but the effect on demand from conversions are found to be rather moderate.

*Sarita Bartlett is a staff scientist at International Energy Studies, Lawrence Berkeley Laboratory, Berkeley, John K. Dagsvik and Øystein Olsen are senior research officers at the Research Department, Central Bureau of Statistics, Oslo. Steinar Strøm is Professor at the Institute of Economics, University of Oslo.

Contents

1 **Introduction and summary** 1

2 **A dynamic discrete-continuous choice model for energy demand** 2
2.1 Introduction 2
2.2 Dynamic discrete choice 4
2.3 The indirect utility function 6
2.4 Short run capacity utilization 7

3 **Estimation and empirical results** 8
3.1 Step 1: Short term capacity utilization 9
3.2 Step 2: Fuel choices in new homes 13
3.3 Step 3: The dynamic Markov model 15

4 **Price effects and elasticity concepts** 18

5 **Concluding remarks** 20

A **Data sources and variable definitions** 22
A.1 General characterisitics 22
A.2 Definitions of variables 22
A.3 Data limitations 24

1 Introduction and summary

In recent years considerable attention has been paid to demand conditions in the market for natural gas and the potential for future growth in gas consumption. These prospects are of particular interest for Norway, being a significant supplier of gas to the European market, and expecting to increase the production in the future. In order to undertake a systematic collection and treatment of information about the energy market and to make forecasts of future gas demand, a formal model is a very useful tool.

This paper presents a formal framework for analysing gas demand for space heating in the residential sector in Western Europe. A ***discrete-continuous dynamic choice model*** is specified. A reasonable interpretation of this approach is that households' decisions on energy consumption are carried out in two steps: In the first step households have to choose between a limited (discrete) number of fuel systems or technologies[1]. Presently, the theoretical model distinguishes between four fuel systems (with indices used in the following in paranthesis): natural gas (1), fuel oil (2), solid fuels (3) and electricity (4). It should be noted that the model thus ignores the existence of 'mixed' technologies. Given this technology choice, a (continuous) decision is made on how intensively the equipment should be utilized. In the short run households can change energy consumption only by varying the intensity in applying the installed equipment, while in a longer time perspective changes in prices and other variables can motivate consumers to convert to another fuel system. Formally, the technology choice involves comparisons of levels of indirect utility attached to the various fuel systems. Consistent with the theory of consumer behaviour, equations describing gas use per household are derived from the indirect utility functions by applying ***Roy's identity***.

The technology decision in the present model is described as a sequence of discrete choices over time, which is an extension of the traditional (static) approach for analysing choice behaviour The formal model is thus dynamic, allowing for ***transitions*** from one technology to another, as opposed to the static MNL model, which is probably best suited for modeling new investment decisions, or choices that are irreversible. Assuming that the unknown (stochastic) terms influencing individuals' utility are independently distributed over fuel alternatives and develop over time according to an extremal process, the decisions can be described as a Markov process. The transition probabilities are functions of the explanatory variables (fuel prices, conversion costs, income level and socio-economic variables) specified in the indirect utility functions, and thus include parameters to be estimated. The expressions for the transition probabilities reveal that this dynamic structure may be interpreted as an extension of the Multi-Nomial Logit (MNL) model to an intertemporal situation.

The dynamic discrete-continuous choice model is estimated from data on energy use in seven European countries. The countries included are Denmark, France, Germany, Italy, the Netherlands, Sweden and the United Kingdom. The data contains energy consumption by end use and type of fuel, figures for existing and new housing stock characterized by choice of fuel system , as well as other 'structural' and economic variables

[1]In this paper the concepts 'fuel system' and 'technology' will be used interchangeably. It should be stressed that the problem of choosing between technologies with ***different efficiences*** is not adressed.

relating to energy use. The data have been constructed from a variety of official and unofficial sources in each country and are organized in a database at Lawrence Berkeley Laboratory, University of California. The variables utilized in the present analysis are defined and briefly discussed in the Appendix I; for a more detailed description of the data, see Schipper et.al. (1985). The model directly applies information on how energy consumption is related to the dwelling stock, dwelling size etc. This distinguishes the present approach from most other studies of energy demand. Only the energy use for *space heating* is modeled. This is by far the most important end use, absorbing more than 70 percent of total energy consumption by households in the countries included in this study. Furthermore, it is probably not quite unrealistic to simplify the analysis and assume that the primary fuel choice is the technology decision for space heating, and that this is done separately from choice of fuel system for water heating and cooking[2].

From the empirical findings of the present analysis we may report the following:

- The short term price elasticity of energy demand is estimated to -0.3 — -0.4.
- Investment choices in new homes are rather price elastic, so that the neoclassical gas price elasticity related to new homes is found to exceed one in absolute value.
- The effect of changes in energy prices on demand through *conversions* is found to be moderate.
- A desicive factor in the natural gas market is the evolution of the distribution network.

With respect to the use of the model framework it should be noted that the demand for housing is not modeled. The dwelling stock must be given exogenously by the model user in order to determine the development of total gas consumption. In the next section the principal features of the formal model and the dynamic discrete-continuous choice approach for analysing energy demand will be presented. Section 3 surveys the estimation procedures and the empirical results, while different types of price responses are discussed and measured by model relations in section 4.

2 A dynamic discrete-continuous choice model for energy demand

2.1 Introduction

In recent years the *discrete choice approach* has gained considerable popularity in the formal modeling of energy demand in the residential sector.[3] As opposed to the traditional

[2]Within a static discrete choice framework Goett and McFadden (1982) undertake very detailed analyses of households' fuel choices in *new buildings*, specifying a Nested Multi-Nomial Logit model, where among many features the choice of hot water alternative is dependent on the initial choice of space heating technology. To estimate such a structure, in particular within the dynamic framework which we are applying, would complicate the model considerably. For example, such a specification should treat installation costs for a water heater as dependent on the type of space heating system chosen 'earlier', and thus either require these data to be available, or leave more parameters to be estimated.

[3]See e.g Dubin and McFadden (1984) and Goett and McFadden (1982)

econometric approach - assuming that energy consumption varies continuously with explanatory variables such as prices and incomes - discrete choice models build explicitly on the fact that an individual consumer usually chooses between a limited number of fuel alternatives. Thus, an important structural feature of households' behaviour may be taken into account in the specification of the theoretical model. Furthermore, as will be discussed below, the discrete variation over choice alternatives causes specific econometric problems when it comes to estimation of the demand model. By utilizing a discrete choice approach these econometric problems are very much brought up to the surface.

In the discrete choice approach the individual behaviour is typically represented by a set of *choice probabilities*. To show the structure of this framework we start with a simple static model for analysing energy demand. Assume an individual consumer who is faced with the problem of choosing between 4 different technologies for space heating. By convention we choose technology 1 to be gas heating. To each technology we define a conditional indirect utility function, V_h, of the following type:

$$V_h = v(z_h) + \epsilon_h \equiv v_h + \epsilon_h \quad , \qquad h = 1, \ldots, 4 \tag{1}$$

V_h is the maximum utility attainable from space heating given that technology h is chosen. V_h consists of a structural part, $v(z_h)$, written short as v_h; and a stochastic term, ϵ_h. In accordance with traditional theory of consumer behaviour, V_h can be regarded as the result of the maximization of a (direct) utility function conditioned by technology h being the most preferable heating choice. z_h is a vector of independent variables or characteristics (i.e. prices, incomes etc.) restricting this maximizing behaviour. Among the characteristics some may be alternative specific, such as the fuel price and the price of the specific type of space heating equipment, while others may be similar for all technologies, e.g. income[4]. The stochastic disturbance term, ϵ_h, is interpreted as expressing the effect of factors on individuals utility that are unobservable to the econometrician, but which are actually taken into account in the optimization process undertaken by the consumers.

The decision criterion for choosing between the available space heating technologies is that alternative h is chosen if this yields a higher utility than any other alternative. The probability for technology h to be chosen can be defined as:

$$\bar{P}_h = Pr(\ V_h = \max_k V_k\ ) = Pr(\ v(z_h) - v(z_k) > \epsilon_k - \epsilon_h\ ,\ \ \forall k \neq h\ ). \tag{2}$$

The explicit form of this probability depends on the form of the indirect utility function and the joint distribution of the stochastic disturbance terms $(\epsilon_k - \epsilon_h)$. By assuming that the unobservable terms are identically and independently extreme value distributed, it can be shown by straightforward integration (see e.g. McFadden (1973)) that the probabilities in (2) can be expressed as

$$\bar{P}_h = \frac{e^{v_h}}{\sum_{k=1}^{4} e^{v_k}} \ , \quad h = 1, \ldots, 4 \tag{3}$$

[4]It should be noted that since the indirect utility is conditioned on a fuel-specific technology, the prices of other fuels or types of equipment do not enter this function

This Multi-Nomial Logit (MNL) model has been frequently applied in empirical discrete choice studies in recent years, due to its simple structure and the fact that effective estimation algorithms have been developed for this model. In the present analysis the MNL structure is also used to formalize households' fuel choices in *new homes*, which constitutes an important part of the complete model. On the other hand, when modeling energy demand in *existing* homes the (static) MNL model is too restrictive. We will argue that in such a case we need a *dynamic* framework, to which we now turn.

2.2 Dynamic discrete choice

In some recent studies of residential energy demand where a discrete choice approach has been applied (Goett and McFadden (1982), Ruderman (1985)) investment decisions in energy using equipment are assumed to be *irreversible*. This means that one has ignored the possibility to convert from one fuel system to another. In a model focusing on long term projections of energy demand the assumption that investments in fuel systems are made "once and for all" seems inadequate. In particular, in the analyses of gas demand, one should include the possibility that in the long run *conversions* from one fuel to another may take place. On the other hand, the process of changing from one fuel to another may obviously be "costly" to a consumer. Both "pure economic" costs and "psychological" factors may be involved when such decisions are taken, both presumably reducing the rate of conversion relative to what would otherwise have been the case.

The purpose of utilizing a *dynamic* discrete choice model, is to capture some of these features which we believe are essential when analysing demand for natural gas[5]. While a static discrete choice model is expressed in terms of state probabilities for the different alternatives, in a dynamic context the behaviour is represented by a set of *transition probabilities*, $P_{sh}(t-1,t)$, defined as

$$P_{sh}(t-1,t) = \Pr(\, V_h(t) = \max_k V_k(t) \mid V_s(t-1) = \max_k V_k(t-1)\,) \tag{4}$$

$P_{sh}(t-1,t)$ is the conditional probability that technology h at time t is chosen given that technology s maximizes utility at time $t-1$.[6].

As in the static model, in order to reach a specific form of the transition probabilities, assumptions have to be made regarding the stochastic structure of the utility process. Our specification builds directly on Dagsvik (1983) and is described in some detail in Dagsvik et.al. (1986). The basic idea is that the households have to choose, not only between the observed energy technologies, they also face alternatives that are *unobservable* to the analysts but known by the decision makers. The extended set of alternatives

[5] We are not aware of any previous study having utilized a dynamic discrete choice model in the analyses of *energy demand*. The general notion of dynamic choice is however discussed in several contributions (see e.g. MacRae (1977) and Heckman (1981)). Oren and Rothkopf (1984) use a dynamic choice framework for analysing market behaviour for new industrial products which bears several similarities to the model proposed in this paper

[6] A more general approach would be to start with the likelihood of a *sequence* of choices and model the joint probability that the optimal technologies at time points $t_1, t_2, \ldots t_m$ are $h_1, h_2, \ldots h_m$. However, since our theoretical model will satisfy the Markov property it is sufficient to specify the one step transition probabilities in (4)

may differ in terms of quality, efficiency and operational convenience. New alternatives arrive according to an (inhomogeneous) Poisson process and they have extreme value distributed utilities. If we assume that the households at each point of time maximizes utility over *all* available choice alternatives, we can then specify an *extremal process* for the utility structure (for a definition see Tiago de Oliveira (1968)). It is intuitively clear that this kind of process implies that the utilities are correlated at different points in time. Dagsvik et. al. (1986) shows that this correlation structure can be expressed as

$$corr(V(t), V(s)) = J(exp(-(t-s)\gamma)), \quad t > s \tag{5}$$

where $V(t)$ is the *unconditional* indirect utility, γ is a nonnegative parameter, $V(t) = \max_h V_h(t)$, and $J(x)$ is the autocorrelation function

$$J(x) = -\frac{6}{\pi}\int_0^x \frac{\log v}{1-v} dv \ , \ \ J(0) = 0, \ J(1) = 1. \tag{6}$$

From 5 and the definition of $J(x)$ given in (6) we observe that when $\gamma = 0$ then $corr(V(t), V(s)) = 1$; that is perfect autocorrelation prevails. When γ increases autocorrelation decreases, and as γ approaches infinity, autocorrelation approaches zero. It may be noted that the specified autocorrelation structure does not depend on the structural part of the indirect utility.

Provided there are no conversion costs the intertemporal choices of heating technologies is a discrete dynamic choice process of the type stated in Dagsvik (1983). By a slight modification of the results therein it follows that this process is a Markov chain with transition probabilities

$$P_{sh}(t-1,t) = \bar{P}_h(t) - \bar{P}_h(t-1)\ e^{-\gamma} \ , \ \ s \neq h \ , \gamma \geq 0 \tag{7}$$

$$P_{ss}(t-1,t) = \bar{P}_s(t) - \bar{P}_s(t-1)\ e^{-\gamma} + e^{-\gamma} \tag{8}$$

where $\bar{P}_k(t)$ is given by (3) above and is the probability of being in state k at time t. When conversion costs are included the choice process is obviously still Markovian, but the transition probabilities now take the form

$$P_{sh}(t-1,t) = Q_{sh}(t) - Q_{sh}(t-1)\ e^{-\gamma} \ , \ \ s \neq h \ , \gamma >= 0 \tag{9}$$

$$P_{ss}(t-1,t) = Q_{ss}(t) - Q_{ss}(t-1)\ e^{-\gamma} + e^{-\gamma} \tag{10}$$

where

$$Q_{sh}(t) = \frac{e^{v_{sh}(t)}}{\sum_{k \neq s} e^{v_{sk}(t)} + e^{v_s(t)}} \tag{11}$$

and $v_{sh}(t)$ is the mean utility of state h given that state s was occupied in period t-1. $v_{sh}(t)$ depends on s when $s \neq h$ due to *conversion costs*. Note that $Q_{sh}(t)$ does not have the same interpretation as $\bar{P}_h(t)$.

The parameter γ reflects the possible stability in the unmeasurable factors influencing utility. The effect of this term may be called *habit persistence*. From (6) we observe that if autocorrelation is zero, (γ is infinitely large), then habit persistence vanishes and the transition probabilities degenerate to the state probabilities. This means that there is no temporal stability in the unmeasurable factors influencing individual decisions on conversions. The opposite situation emerges when $\gamma = 0$, which is equivalent to perfect autocorrelation. From (7) we observe that when $\gamma = 0$, $P_{sh}(t-1,t) = \bar{P}_h(t) - \bar{P}_h(t-1)$.

Obviously, changes in 'observable' choice conditions, such as e.g. prices, may tend to make an alternative technology more attractive than it was one period earlier, and motivate the consumers to convert. When deciding upon this, conversion costs will also have to be taken into account. The existence of state dependent conversion costs may be said to reflect *structural state dependence* in the behavioural model. We may interpret structural state dependence as the effect of past choices actually experienced by the households, while habit persistence represents the effect of past choices from the observers viewpoint.

2.3 The indirect utility function

The functional form of the indirect utility function should of course be consistent with basic properties derived from consumer behaviour. Apart from that, our choice of functional form has primarily been motivated by the need for specifying a rather simple structure which e.g. is manageable when coming to estimation of the model. In Hanemann (1984) several specific indirect utility functions which may be suitable as representations of discrete/continuous consumer choices are suggested . Our choice is a modification of one of these utility models. More precisely, the following state dependent indirect utility function has been chosen:

$$v_{sh} = \frac{\alpha_0}{\alpha_1 - 1} p_h^{1-\alpha_1} + \alpha_4 + \alpha_2 r_h - \frac{1}{\alpha_3} e^{-\alpha_3 y + \sum_{k=2}^{6} \beta_k z_k} + \beta_1 (1 - \delta_{h1}) z_1 - (1 - \delta_{sh}) c_s \quad (12)$$

where $\alpha_0 > 0,\ \alpha_1 \neq 1, \alpha_3 \neq 0$

The variables in this relation are defined and discussed in some detail in Appendix A. In short, the symbols in (12) have the following interpretation:

p_h is the real price of fuel h per energy unit.

r_h is the user cost of energy using equipment of type h relative to the consumer price index. Geometric depreciation is assumed.

y is real disposable income per household.

z_1 denotes the gas network saturation.

z_2 is the relative penetration of district heating.

z_3 is the relative share of single family dwellings.

z_4 is the average dwelling area.

z_5 is a climate factor, measured as the total number of heating degree days.

z_6 denotes the relative penetration of central heating systems.

δ_{h1} is a dummy variable related to whether the household is connected to the gas network or not; $\delta_{h1} = 1$ if $h = 1$ (i.e. in the case of natural gas), $\delta_{h1} = 0$ if $h \neq 1$ (thus, if $\beta < 0$ then an expansion of the gas network (increase in z_1) has a positive impact on the utility).

δ_{sh} is a dummy related to whether a conversion takes place or not; $\delta_{sh} = 1$ if $s = h$, $\delta_{sh} = 0$ if $s \neq h$.

c_s represents conversion costs (measured as loss in utility) of changing fuel system.[7]

The specification (12) ensures that the indirect utility is a decreasing and convex function in prices. Moreover, the conditional indirect utilities increase with increases in income. The additivity in the terms involving price of energy and user costs of capital respectively, implies that energy prices affect the investment and transition decisions (through the state probabilities) as well as decisions on capacity utilization. The user cost of capital affects investment and transition decisions only, since this variable vanishes when the conditional demand functions are derived (see below).

A general question regarding the explanatory variables in the indirect utility function is whether they should refer to current or expected values. Since investment choices in various types of equipment are involved, the latter interpretation is preferable, and an explicit expectation mechanism should be introduced, at least for some of the independent variables. However, so far we have not attempted to specify any explicit expectation structure.

2.4 Short run capacity utilization

As mentioned in the introduction, the present model assumes that when consumers have chosen a specific technology, they can vary the intensity or utilization rate of the heating- equipment continuously. At this stage we can apply the theorem known in the literature as ***Roy's identity***. This states that the (uncompensated) demand functions for the various commodities can be derived from a fully specified indirect utility function simply by differentiation. If we introduce the notation

x_h as consumption of fuel h per household/dwelling,

Roy's identity states that the demand functions for fuel h can be derived from the indirect utility function as

[7] In Dagsvik et.al. (1986) conversion cost was specified as a reduction in *income*. The specification of c_s in (12) should be regarded as a normalization of the pure economic costs and is introduced for the sake of simplifying the econometric model.

$$x_h = -\frac{\partial\, v_{sh}/\partial\, p_h}{\partial\, v_{sh}/\partial\, y} \tag{13}$$

Relation (13) illustrates an essential feature of the discrete/continuous choice model: having solved the technology choice problem, demand relations describing the utilization of a given equipment follows directly by applying equation (13)[8].

From the chosen form of the indirect utility function (12) we can now derive the following set of demand equations[9]:

$$\log x_h = \log \alpha_0 - \alpha_1 \log p_h + \alpha_3 y - \sum_{k=2}^{6} \beta_k z_k \tag{14}$$

3 Estimation and empirical results

The international energy data used for estimating the model (cf. section 1 or the Appendix A) are not strictly 'individual' data. What we actually have are *country averages* for the various variables in the model. The data may thus be regarded as what in the literature is called *grouped* data (Maddala (1983)). Another principle limitation set by the available data is that the LBL database at present does not provide sufficient information on conversions from one technology to another. If such data on observed transitions had been available, the estimation procedure would have been simplified, as this would have allowed us to employ directly the relations for transition probabilities for calculating coefficients of the model.

MacRae (1977) discusses several procedures for estimating a Markov model with time dependent transition probabilities in a situation where only the state frequencies at different points of time - not the transitions - are observable. Following her results, we can specify the following set of stochastic equations:

$$\hat{P}_h(t) = \sum_{s=1}^{4} \hat{P}_h(t-1) P_{sh}(t-1,t) + \mu_h(t)\ , \quad h = 1 \ldots 4 \tag{15}$$

where $\hat{P}_h(t)$ is the observed frequency of technology h. As shown in MacRae (1977), the means of the residual terms, $\mu_h(t)$, in (15) equal zero. The covariance matrix - conditional on the states occupied at time $t-1$ - can also be readily calculated, revealing that the equations in (15) have heteroscedastic disturbances. Thus, GLS is a preferable procedure for estimating the parameters of the model.

All parameters of our theoretical model may be identified in the set of Markov-equations in (15). To base estimation of the coefficients solely on these dynamic relations would, however, imply that not all available information about the structural model is

[8] A detailed presentation of the interactions between a discrete choice problem and the (conditional) decisions on capacity utilization is given by Hanemann (1984)

[9] Strictly, by applying Roy's identity directly on (12) the derived demand are *conditioned* on state s. However, since these actually are independent of s, the unconditional demand, x_h is easily arrived at by averaging (see Dagsvik et.al. (1986))

utilized. The information on fuel choices in new homes and capacity utilization given type of equipment (the demand equations derived by Roy's identity) should be utilized as well. However, for computational convenience we have carried out the estimation procedure in three stages:

> **Step 1:** In the first step the demand equations derived from *Roy's identity* are used to estimate the parameters which can be identified from these relations.
>
> **Step 2:** The next step in the estimation procedure is to utilize the data on *new constructions*, insert the parameter estimates identified in step 1 and use a conventional static MNL model to estimate initial fuel choices.
>
> **Step 3:** In the final step, the dynamic Markov model is used to estimate the remaining parameters of the model, which actually can be restricted to c_s and γ, i.e., the central coefficients related to the dynamic features of the model. Again, all parameter estimates obtained previously are substituted into these regressions.

This sequential estimation procedure is obviously not optimal from a strict econometric point of view. The most efficient estimation method would have been to calculate all parameters *simultaneously*, utilizing all information contained in the three stages at the same time. However, with all non-linearities and other complex features of the specified theoretical model, this would be a quite complicated task. The stepwise procedure outlined above should thus be regarded as a compromise between what is desirable on theoretical grounds and what can be implemented. The actual estimation carried out in the various stages and the corresponding empirical results are described below.

3.1 Step 1: Short term capacity utilization

A first point to be noted is that the random variables specified in the conditional indirect utility functions cancel out when applying Roy's identity. Thus, there are no interactions between the random utility model and the derived demand functions. With respect to the stochastic specification in the utility functions, the demand functions are deterministic.[10] However, we now introduce the (reasonable) additional assumption that there are *measurement errors* involved in the observed variables included in the demand functions. This means that the following stochastic specification of (14) is assumed:

$$\log x_h = \log \alpha_0 - \alpha_1 \log p_h + \alpha_3 y - \sum_{k=2}^{6} \beta_k z_k + \eta_h \tag{16}$$

η_h is a stochastic error term interpreted to represent meaurement errors, with zero mean and constant variance. Furthermore, the demand relations for the different fuels are assumed to be uncorrelated. This makes Ordinary Least Square (OLS) a reasonable

[10] With a more general formulation of the stochastic terms in the utility function, the derived demand functions might alternatively have been stochastic, i.e. included the ϵ-terms. Such interactions between probability structure in the discrete choice problem and the conditional demand structure are discussed in detail in Hanemann (1984).

estimation method. Since parameters are identical over fuel types this is done efficiently by using all four demand equations in (16) simultaneously in the regression. It may be noticed that a majority of the parameters in the model as a whole is actually identified in the present stage of the estimation procedure.

The estimation results are presented in table 1. The various numbered models refer to a test procedure visualized in figure 1 below. Looking first at the figures in the first row, i.e. the calculation results with no a priori restrictions imposed, Model I, the following features may be stressed:

- Most parameter estimates have the expected sign. In particular this applies to the coefficients related to energy prices and incomes, implying negative energy price elasticity and postive income elasticity.

- Five out of eight coefficents are significantly estimated at a 5 percent significance level. From an economic point of view it is interesting to note that among these are the energy price coefficient (α_1) and the income parameter (α_3). But also the parameters expressing the effects of district heating (β_2) and the climate (β_5) have rather small variances estimated.

- The *energy price elasticity*, directly expressed by the negative of α_1, is estimated to -0.42. This should be interpreted as a *short term* price elasticity, since it is defined conditional on that a specific fuel is already chosen. The figure is in reasonable accordance with estimates obtained in empirical studies previously undertaken (see e.g. Bohi (1981)), perhaps somewhat on the higher side (in absolute value).

- The estimate of the income parameter, α_3, as is the case for all the β-parameters, is dependent on the unit of measurement of the variables involved. The expression for the short term *income elasticity* is given by

$$E_y^s \equiv \frac{\partial \log x}{\partial \log y} = \alpha_3 y \tag{17}$$

  For the average income level in the sample, which is approximately 15000 US dollars, this implies an income elasticity of 0.75.

  It is interesting to focus somewhat more closely on this estimate of the income elasticity. Again, it is in reasonable coherence with figures from several previous energy demand studies, and not very far from what one would expect as a measure of a short term income effect on energy consumption. However, in a number of studies focusing particularly on the demand for natural gas in Western Europe there have been severe problems of obtaining reasonable estimates of the income elasticity. Typically the estimates obtained have been far too high in magnitude. The reason for this common problem is the rapid evolution of the European gas market and the rapid increase in gas consumption experienced through the last 15 years. Traditional econometric models using some sort of ad hoc lag structure to distinguish between short- and long term effects have not been able to explain the

Model	Parameter estimates (t-values) α_0	α_1	α_3	β_2	β_3	β_4	β_5	β_6	R^2	DW	SSR
I	2.64	.42	.005	-1.11	-.06	-0.01	-.0002	0.35	0.39	2.15	30.309
	(5.19)	(7.80)	(2.10)	(-2.18)	(-0.13)	(-1.25)	(-3.05)	(0.90)			
II	3.00	.42	.006	-1.08	-.44		-.0002	0.14	0.38	2.14	30.599
	(7.27)	(7.81)	(2.65)	(-2.12)	(-1.22)		(-3.48)	(0.40)			
IIIa	3.04	.42	.004		-0.58		-.0002	-0.44	0.37	2.07	31.445
	(7.29)	(7.68)	(1.91)		(-1.62)		(-3.12)	(-2.03)			
IIIb	3.03	.42	.005	-0.92	-0.48		-.0002		0.38	2.13	30.629
	(7.49)	(7.87)	(3.05)	(-2.93)	(-1.41)		(-3.82)				
IV	2.84	.40	.006		-0.39		-.0003		0.35	2.04	32.237
	(6.94)	(7.41)	(3.29)		(-1.07)		(5.43)				

Table 1: Estimation of step 1

rapid expansion of the gas market; strong increases in gas demand together with a moderate income growth have resulted in that the income elasticity has been overestimated. A preferable feature of our discrete-continuous choice model is that the income effect is identified in households *which have already chosen their heating equipment*. The estimate of the income parameter is therefore not influenced by the fact that the number of gas customers has increased strongly through the estimation period. In our theoretical model the latter feature should under ideal circumstances be represented by the development of the gas distribution network (z_1).

- It may be noticed that the effect of *district heating* on energy consumption of other fuels is estimated to be *positive* (the negative of β_2). This may seem a bit odd, since one immediately should think that increased presence of district heating should lower the need for other fuels. However, one should remember that the model assumes away dual fuels systems in the households. Accordingly, the obtained result simply indicate a positive correlation between energy (unit) consumption and the penetration of district heating between countries and over time. On this background it may of course be questioned whether it makes much sence to specify district heating as an argument in a short term demand equation. Therefore, we have experimented with excluding this variable from the model.
- The *climate variable*, z_5, is measured as the number of (heating) degree days during the year in the various countries. The estimated figure implies a 'climate elasticity'

$(\frac{\partial \log x}{\partial \log z_5})$ of 0.62.

- Unit energy consumption is positively correlated with the *share of single family dwellings*, as one would expect, but the estimate obtained is very uncertain.
- Increases in the *average dwelling area* also tend to raise unit consumption. This variable is probably strongly correlated with the income level, and we have therefore reestimated the model exclusive of this variable to see whether the income elasticity is changed.

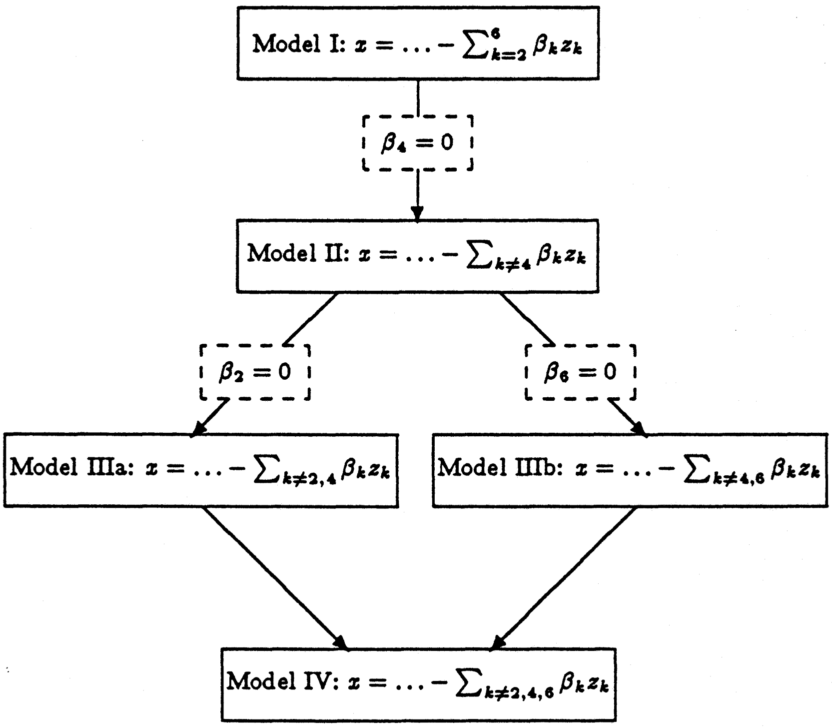

Figure 1: Test scheme for the z-variables

We have also carried out a number of additional estimation with a priori *restrictions* on the coefficients. More specifically, we have sucessively excluded some of the z-variables and tested nested hypothesis against each other. The test procedure is based on Hendry (1974), utilizing hyphotesis about the distribution of proportions between related sum of squared residuals (SSR's)[11]. An overview of the different models is given in figur 1,

[11]Define the statistics $Z = T \log (\frac{SSR_j}{SSR_i})$ where T is the number of observations, and let v_i and v_j denote the number of parameters in model i and j respectivly ($v_i > v_j$). Asympthotically we then have: $Z \sim \chi^2_{v_i - v_j}$

and the corresponding empirical results are also included in table 1.

In model II we have excluded the variable z_4, the averge dwelling area, as this is assumed to be highly correlated with income. The estimate of the income parameter, α_3, is not significantly changed by this restriction. All other models are then nested to II. The test procedure supports the hypothesis that $\beta_4 = \beta_6 = 0$, while the models IIIa and IV are rejected.

3.2 Step 2: Fuel choices in new homes

In the decisions on fuel technologies in new homes, obviously no dynamic elements are involved and the behaviour may be described by a conventional static discrete choice model. It is easily seen that when both conversion costs and 'habit persistence'are excluded from the theoretical model outlined in section 2, the state probabilities degenerate to the traditional MNL form, (3). Furthermore, it seems reasonable to assume that the same structure for indirect utility prevails when describing this behaviour. The structure of indirect utility for investing in new equipment is thus given by (12) except that the conversion costs c_s in this case is excluded.

As already mentioned, the general procedure both for step 2 and 3 is to substitute parameter estimates from the previous steps into the present regressions. The results presented below are based on the unrestricted model in step 1. Actually, all of these coefficients cancel out except the price elasticity α_1. In the following it is convenient to define the following auxilliary variables: $ap_h = (1/(\alpha_1 - 1))p_h^{1-\alpha_1}$, where α_1 is fixed to 0.42 (cf. table 1).

As mentioned in the introduction, *individual* observations of fuel choices by households are not available. The data used in the present step (and also in step 3) consist of *shares* in each country/year utilizing the various fuels for space heating. These empirical frequencies are interpreted as observations of the corresponding theoretical choice probabilities. An equivalent representation of the model (3) is given by

$$\log(\frac{\bar{P}_h}{\bar{P}_1}) = v_h - v_1 \ , \quad h = 2,3,4 \tag{18}$$

By replacing the theoretical choice probabilities in (18) with observed frequencies, $\hat{\bar{P}}_h$, we obtain

$$\log(\frac{\hat{\bar{P}}_h}{\hat{\bar{P}}_1}) = v_h - v_1 + \varphi_h \ , \quad h = 2,3,4 \tag{19}$$

where φ_h is a stochastic error term defined by (19). From the 'multinomial structure' of the choice problem facing the individual households it is easily seen that $E(\varphi_h) = 0$, and it can furthermore be shown that the variances and covariances are of the following form:

$$var\ \varphi_{hi} = \frac{1}{n_i}(\frac{1}{\bar{P}_{hi}} + \frac{1}{\bar{P}_{1i}}) \tag{20}$$

Model	Parameter estimates (t-values)		
	α_0	α_2	β_1
OLS	-0.14 (-1.20)	0.89 (3.27)	-2.17 (-2.55)
ZELLNER	-0.007 (-0.07)	0.47 (1.84)	-2.24 (-2.87)
ZELLNER, S-iterativ.	0.33 (3.82)	-0.14 (-0.70)	-4.12 (-7.45)

Table 2: Estimation of step 2

$$covar\ (\varphi_{hi}, \varphi_{si}) = \frac{1}{n_i \bar{P}_{1i}}\ , \qquad s \neq h \tag{21}$$

The estimation procedure that may be suggested by relation (19) is the *minimum chi-square MIN* X^2 -method, described e.g. in Amemiya (1981). This estimator is obtained by applying a weighted least square procedure to (19). However, in our case n_i is very large so that we have $var\ \varphi_{si} \approx 0$.

Since the explanatory variables are highly aggregated, it seems unreasonable to expect the coefficients in the structural part of the utility function to be constant over time. We therefore assume that the coefficients may be random. This specification implies additive random disturbances in the mean utility function v_h. These random terms are supposed to account for unobserved price variation across consumers as well as measurement errors in prices and other independent. We have assumed unspecified correlation structure between these disturbances. With only one endogenous variable on the left hand side of each equation, as in (19), *Zellner's seemingly unrelated regression method* is convenient. This procedure is applied both with and without an iterative process on the 'S-matrix' (the variance-covariance matrix of the residuals). In addition, we have also experimented with using OLS, again utilizing all (three) equations in (19) simultaneously.

By inserting the expressions for the indirect utility function into (19), and also using the auxilliary variables introduced above, we obtain the following econometric specification of the model:

$$\log(\frac{\hat{\bar{P}}_h}{\hat{\bar{P}}_1}) = \alpha_0(ap_h - ap_1) + \alpha_2(r_h - r_1) + \beta_1 z_1 + \xi_h\ , \quad h = 2,3,4 \tag{22}$$

From this equation it is revealed that in this step the parameters α_2 (expressing the impact on fuel choices by changes in prices on heating equipment) and β_1 (measuring the effect of expanding the gas distribution network) are identified. The estimation results based on equation (22) with different estimation procedures applied are given in table 2.

Based on arguments from economic theory one should expect both α_2 and β_1 to be negative; an increase in the price of energy using equipment should cause a negativ shift in the conditional indirect utility, and increased gas network availability must be expected to imply an increase in the utility related to natural gas. Looking at the figures in table 2 we see that β_1 is significantly estimated and has the correct sign in all the presented models. α_2 is estimated to be positive in the first two models, while the estimate has the correct sign in the S-iterative Zellner estimation, but with a large standard deviation. We may thus conclude that the model and the data available have not given us any reliable estimate of how changes in heating equipment prices affect fuel choices. This is not very surprising, since the quality of the data on equipment prices is rather poor.

When passing over to step 3, again substituting parameter estimates from both preceeding sections, we have chosen to utilize the results from the S-iterativ Zellner model.

3.3 Step 3: The dynamic Markov model

In this final step, the coefficients related to the dynamic features of the model, the correlation parameter γ and the conversion costs, c_s, are estimated within the Markov model (15). Let us repeat the set of relations for the transition probabilities

$$P_{sh}(t-1,t) = Q_{sh}(t) - Q_{sh}(t-1)\, e^{-\gamma} \qquad s \neq h \ , \tag{23}$$

and define the probability of being in state h (according to the Markov model) as

$$P_h(t) = \sum_{s=1}^{4} P_s(t-1) P_{sh}(t-1,t) \ . \tag{24}$$

Here we should remind ourselves of the fact that the P_{sh}- variables in (23) are expressed in terms of the conditional indirect utility function, i.e. they depend also on the specified conversion costs. This implies that the econometric model which is obtained by substituting (23) into (24) becomes very complicated. Remember also from section 2 that in the special case of no conversion costs the P_{sh}-variables degenerate to MNL choice probabilities, given by relation (3). Note then that the structure of these relations now are ***completely known*** by the results from step 1 and 2. This feature will be utilized in the following to simplify the econometric specification in the present stage.

The simplification is obtained by introducing a transformation of the conversion cost variables. Let us define:

$$d_s = \frac{e^{v_s(t)}}{\sum_{k=1}^{4} e^{v_k(t)}}(e^{c_s} - 1) \ = \bar{P}_s(t)(e^{c_s} - 1) \tag{25}$$

Treating d_s as a constant, as it will in the following, c_s varies 'proportionally' with the inverse of the choice probability, $\bar{P}_s(t)$. If the probability of choosing a specific fuel is 'high', the corresponding conversion cost is relatively low, and vice versa.

It is now easily shown that the transition probabilities can be written in the following way:

$$P_{sh}(t-1,t) = \frac{1}{1+d_s}[\bar{P}_h(t) - \bar{P}_h(t-1)\, e^{-\gamma}] \qquad s \neq h\ . \tag{26}$$

By transforming the conversion costs we have thus expressed the transition probabilities in terms of the MNL state probabilities, $\bar{P}_h(t)$. The term between the brackets is the transition probability in the case where there are no conversion costs. Let in the following this be denoted by $\bar{P}_{sh}$, and define furthermore

$$a_s = \frac{1}{1+d_s} \tag{27}$$

We can then summarize the modified econometric model as follows:

$$\bar{P}_{sh}(t-1,t) = \bar{P}_h(t) - \bar{P}_h(t-1)e^{-\gamma} \quad , s \neq h \tag{28}$$

$$\bar{P}_{ss}(t-1,t) = \bar{P}_s(t) - \bar{P}_s(t-1)e^{-\gamma} + e^{-\gamma} \tag{29}$$

$$P_{sh}(t-1,t) = a_s \bar{P}_{sh}(t-1,t) \quad s \neq h \tag{30}$$

$$P_{ss}(t-1,t) = 1 - \sum_{k \neq s} a_s \bar{P}_{sk}(t-1,t) = 1 - a_s(1 - \bar{P}_{ss}(t-1,t)) \tag{31}$$

$$\hat{P}_h(t) = \sum_{s=1}^{4} \hat{P}_h(t-1) P_{sh}(t-1,t) + \mu_h(t) \tag{32}$$

The expressions $\hat{P}_h(t)$ in (32) are *observed frequencies* of the different fuel systems in existing homes, and the residual terms $\mu_h(t)$ are defined by this relation. Parallell to what was the case in step 2, the multinomial structure of the individual choice model has certain implications for the variance-covariance matrix of of these residuals (see e.g. MacRae (1977)). As explained above, however, we get in our case that $\mu_h \approx 0$. Similarly to step 2 we have based the estimation on general regression procedures - like OLS and Zellner estimates.

So far we have assumed that the costs of converting to another technology differ between fuels. Obviously, the problem is considerably reduced by restricting the *transformed* conversion costs to be *identical*; i.e. $a_1 = \ldots = a_4 = a$. From (28) - (32) we then arrive at the following simple regression model:

$$\hat{P}_h(t) - \hat{P}_h(t-1) = a\,(\bar{P}_h(t) - \hat{P}_h(t-1)) + b\,(\hat{P}_h(t-1) - \bar{P}_h(t-1)) \tag{33}$$

where $b = a\ e^{-\gamma}$. Each of the equations in (33) (only three are included in the estimation due to the restriction that the probabilities sum to unity) is linear in the parameters.

Model, method	Parameters (t-values)						
	a	a_1	a_2	a_3	a_4	b	γ
OLS	0.07 (0.41)					0.055 (0.30)	0.28
ZELLNER	0.21 (1.19)					0.17 (0.92)	0.24
ZELLNER S-iterativ	0.33 (2.00)					0.27 (1.60)	0.20
OLS		-8.66 (4.17)	-0.65 (-2.17)	-0.89 (-1.51)	5.82 (2.27)		0.008 (0.86)
ZELLNER		8.28 (4.3)	-0.69 (-2.24)	-0.06 (-0.11)	4.71 (1.80)		-7.610^{-4} (-0.07)
ZELLNER S-iterativ		9.08 (5.06)	-0.61 (-1.98)	-0.39 (-0.65)	3.83 (1.51)		-0.01 (-0.77)

Table 3: Estimation of step 3

Estimation results both from (33) and from the more general model with no restrictions on the conversion costs are presented in table 3. Also at this point we have experimented with using both OLS and Zellner regression methods.

From the results in table 3 we may notice the following features:

- Due to its interpretation in the theoretical model it is required that $\gamma \geq 0$. Furthermore, from the defintions of d_s and a_s it is seen that $0 \leq a \leq 1$ should be expected. These requirements are fulfilled in all the calculations based on a unique conversion cost over fuels. The most significant parameter estimates are obtained in the iterative Zellner procedure.

- We notice that the estimate of the correlation parameter γ is rather low in magnitude (actually it is not significantly different from zero in any of the alternative regressions. In other words: the model indicates a high positive autocorrelation between the stochastic elements influencing peoples' fuel choices, implying that there is a set of unknown factors motivating households to stick to a fuel system once chosen.

- The results from the regressions with no restrictions on the a-coefficients imposed are observed to be in conflict with the a priori restrictions; with a_2 and a_3 estimated to be negative (implying negative transition probabilities) and a_1 and a_4 estimated to exceed one (implying $d_1,\ d_4 \leq 0$).

Model	a	d	Conversion costs				
			c_1	c_2	c_3	c_4	γ
ZELLNER, S-iterativ	0.33	2.03	2.42	1.60	2.45	3.11	0.20

Table 4: Conversion costs

From the estimates of a, the transformed conversion costs, d and c_s can be readily calculated as $d = (1-a)/a,\ \ c_s = \log((d+\bar{P}_s)/\bar{P}_s)$. The values for these parameters corresponding to the Zellner S-iterativ estimation of model (33), and calculated for sample means of the $\bar{P}_s$ - variables, are reported in table 4.

4 Price effects and elasticity concepts

Based on the empirical results surveyed in the previous section we now turn to describing how the model traces the impacts on energy demand caused by changes in energy prices. Estimates of such price effects may be interesting per se, and in a process of using the model to analyze feasible ranges of future demand. The discussion is carried out in terms of price elasticities, and the focus is on the demand for natural gas.

Due to its dynamic structure and the distinction between technology choices and capacity utilization several kinds of elasticity concepts may be defined within the present model. The **short run effect** on the demand for natural gas is given directly by the parameter α_1, as explained in section 3. This follows from the logic of the model; in the very short run the consumers can vary their energy consumption only by changing the intensity of use of the given equipment. Thus, denoting the short run elasticity by E_S, we define

$$E_S = \frac{\partial \log x_1}{\partial \log p_1} = -\alpha_1 \tag{34}$$

In a longer time perspective changes in prices may influence consumers choices of fuel system in new homes and also motivate existing households to convert to another heating technology. The elasticity of the probability of choosing gas among *new dwellings* with respect to the gas price, E_N, is

$$E_N(t) = \frac{\partial \log \bar{P}_1^n(t)}{\partial \log p_1(t)} = -\alpha_0 p_1(t)^{1-\alpha_1}(1 - \bar{P}_1^n(t)) \tag{35}$$

where the notation $\bar{P}_1^n$ is introduced at this point to distinguish this variable from the state probability for the *existing* dwellings, P_1^e, in the Markov equation below. While the short run elasticity measured by α_1 is a constant, the price effect on the choice probability in (35) varies over time. The elasticity of the expected gas consumption in new homes is the elasticity of $\bar{P}_1^n x_1$ with respect to the price of gas, which of course is equal to $E_S + E_N$. We term this the *long run neoclassical elasticity* of gas demand.[12]

A third type of price effect on energy demand implied by the present model comes through *conversions* to or from other fuel systems. Principally, a change in e.g. the gas price influences all transition probabilities at each point of time in a specific scenario. To arrive at formal expressions for this kind of energy demand impacts is complicated, and the easiest way of obtaining empirical characteristics is to undertake simulations and sensitivity analysis on the estimated model system.

The impacts from conversions affecting gas consumption are summarized in the effect on the state probability, $P_1^e(t)$. In order to briefly discuss the nature of this effect, it may be useful to repeat the expression for this probability from the Markov model, i.e.

$$P_1^e(t) = \sum_{s=1}^{4} P_s^e(t-1) P_{s1}(t-1,t) \tag{36}$$

Starting out from a given set of state probabilities in period t-1, $P_s^e(t-1)$, the *first year* conversion effect for gas stems from changes in the transition probabilities, $P_{s1}(t-1,t)$, in (36). This effect may be called the **intermediate term conversion effect**.

With respect to **long term conversions effects**, it is decisive whether the price change is temporary or permanent. If the price change only relates to time t_0, and if furthermore expectations are static, only transition probabilities for period t_0 and $t_0 + 1$ are affected[13]. If on the other hand a price change is permanent, transition probabilities in all future periods are affected, and thus reinforcing a tendency to convert to or from another fuel. Denoting the **long run conversion elasticity** by E_C, the long run elasticity of expected gas demand in existing dwellings, i.e. taking conversions into account, is the elasticity of $P_1^e x_1$, or $E_S + E_C$

In order to illustrate the magnitude of the different types of price effects, we have calculated a set of price elasticities of the various kinds based on data for *France*. The results are presented in table 5. The estimates of the conversion effects are obtained by simulating the empirical model from 1984 to 2000. A reference scenario is constructed by making rough estimates for the exogenous variables, and thereafter the price path of natural gas is given a permanent 10 percent increase in 1985.[14]

It is interesting to notice from table 5 that the estimated results imply a considerable flexibility and price sensitivity in investment decisions in new constructions. The relative demand impact through fuel choices in new homes of an increase in the gas price is twice the magnitude of the short run effect. Together, these two effects imply an neoclassical gas demand elasticity of 1.2 - 1.3 in absolute value. On the other hand, consistent with

[12]The term "neoclassical" refers to the feature that the demand respons to changes in prices is not restricted by costs of conversion or refers to the long run demand respons if γ were infinitly large.

[13]It is important to notice, however, that state probabilities for later periods in general are influenced even by a temporary price change

[14]Since the model is non-linear, *total elasticities* of this kind may depend on the reference scenario chosen. Experiences indicate, however, that this dependence is rather weak

Elasticities	1985	1990	2000
Unit consumption, short run (E_S):	-0.420	-0.420	-0.420
Choice prob., new homes (E_N):	-0.819	-0.843	-0.886
Long run neoclassical ($E_S + E_N$):	-1.239	-1.263	-1.306
Conversion effect, 1. year (E_{C1}):	-0.052		
Long run conversion effect (E_C):		-0.070	-0.083
Long run conversion elasticity ($E_S + E_C$):		-0.510	- 0.523

Table 5: Estimated price elasticities for France

the high 'habit persistence estimated (cf. the previous section), the changes in the share of gas customers due to conversions are very moderate; the elasticity, however, increasing over time. We observe that the long run conversion elasticity in expected gas demand in existing dwellings is substantially lower in absolute value than the long run neoclassical elasticity.

5 Concluding remarks

Traditionally, residential energy demand modeling has pursued two different lines of development. The first is the *econometric approach*, where demand relations are derived from 'smooth' economic consumer behaviour. The second is the *engineering approach*, which utilizes detailed information of market penetration of different types of energy using equipment and their energy intensity. The discrete-continuous choice approach integrates elements from both these procedures of energy demand modeling. It explicitly builds on the observation that private households choose between a limited number of fuel technologies. Furthermore, it links energy consumption to characteristics of the dwelling stock in a way that makes it easier to take into account structural barriers and limitations on energy consumption. At the same time the market penetration of the different technologies and their intensity of use are based on relations drawn from economic theory, stressing the impact of prices and incomes for the evolution of energy demand.

To harvest the full advantages of using a discrete-continuous choice approach requires that *micro data* are available. This points at an obvious weakness of the empirical results presented in the this paper: only aggregate data were available, restricting both the formal specification and the estimation of the model. As discussed in some detail in the Appendix A, other problems in the model specification can also partly be related to the present data situation (the abscence of district heating as a separate fuel, the lack of choice of central/non-central heating etc.). In spite of these limitations, the estimation of the model has yielded some empirical results which are rather resonable. In particular, the estimates of the price- and income effects seem fairly robust, and it is also interesting to notice the significance of the distribution network on the penetration of natural gas. With respect to the dynamic features of the model, the estimation traced a significant degree of habit persistence by the consumers, i.e. 'reluctance' to convert to another heating technology. However, in a long term perspective the effects of conversions on energy use forecasted by the model will not be negligible.

A main purpose of constructing a model as described in this paper is to establish a framework for projecting residential gas demand in Western Europe. How the model framework is utilized for this purpose is described in Dagsvik et.al. (1986). A basic feature of the model is that in order to estimate future demand for natural gas one has to make assessments of the evolution of the dwelling stock in the various countries. This of course requires extra efforts by the model user, but as mentioned above we tend to regard the explicit connection between energy use and the dwelling stock as a strong side of the model rather than a weakness. Given this input and assumptions on the other exogenous variables, the specified model then calculates

- Fuel choices in new homes (the static MNL model).
- The penetration of different fuel systems in existing homes (the dynamic Markov model).
- Energy consumption per dwelling (short term demand equations).
- Total consumption of natural gas by country and in Western Europe (by aggregation).

With respect to future work it obviously would be of considerable interest to utilize a similar type of model framework on a set of truly individual data. Furthermore, since the model involves investment decisions, the impacts of *expectations* should be given specific attention, most preferably within a more explicit dynamic optimization framework. Finally, more efforts - both academic and empirical - will be needed to make the model capture the existence of secondary heating, and also to include energy demand for other end uses than space heating.

A Data sources and variable definitions

A.1 General characterisitics

The data utilized in the estimations presented in this paper are from Lawrence Berkeley Laboratory's (LBL's) OECD residential energy use data base. These data are collected from a large number of official and private sources in each country and year. Consumption data by end use and fuel type have been constructed by using "bottom up" information on dwelling stock and measurements of unit energy consumption. For major OECD countries economic data and a number of structural indicators relevant for analyzing energy demand are also included in the data base. Considerable efforts have been put to achieve consistent definitions of variables to allow for comparisons between countries. For most countries the time series begin in the early 1960's and extend through 1983. It should be stressed that presently the data base does not contain continuous time series for the different countries.

The countries and years included in the data sample are shown in the following table:

country	*years*
Denmark	1965,70,72,77,80,81,82
France	1962,1973-81
Germany	1960,65,70,72,75,78,80,81,82
Italy	1978,80
Netherlands	1973,78,81,83
Sweden	1963,65,70,72,75,78,80,81,82
United Kingdom	1970,75,78

A.2 Definitions of variables

In section 2 the variables included in the formal model were defined very briefly. Below the variables and the data used are listed and discussed in some more detail.

Space heating fuels

In the model four different fuels are assumed to be available for space heating use: ***fuel oil***, ***natural gas***, ***solids*** and ***electricity***. Fuel oil includes LPG in all countries, except Germany, where it was included in natural gas in the primary data sources.

Dwellings by fuel type

The LBL data base has collected information of number of dwellings with different fuel types. Both figures for total stocks and new constructions are available, but for the latter some observation points are missing compared to what is indicated in the table above. The dwelling figures provide observations of the frequencies which are empirical counterparts to the choice probabilities, $P_h(t)$ and $\bar{P}_h(t)$, in the formal model.

Intensity of use

The variable $x_h(t)$ denotes consumption of fuel h per household (dwelling). Energy consumption is measured in GJ and defined as delivered amount of energy (type h) used for space heating purposes.

Prices and income data

Prices and income data are in 1981 US dollars. The current local prices were deflated using consumer price indices of the individual countries, with 1981 as the base year. The real prices and incomes were then converted to dollars using the appropriate exchange rates for the same year. *Fuel prices*, p_h, are average residential fuel prices per GJ including taxes. The price of solid fuels is assumed equal to the price of coal, since proper wood prices were difficult to obtain. *Income per household*, y, is defined as national disposable income per household. *Equipment prices*, q_h, are measured as average equipment prices for the different fuels. The data situation for these variables is weak, and the utilized figures are constructed from scattered estimates from a few countries and years.

District heating penetration

The variable z_2 is defined as the share of dwellings using district heating for space heating. Strictly this technology should have been included as a separate 'fuel' in the model, but prices and other data were not available.

The share of Single Family Dwellings

The share of single familiy dwellings, z_4, is defined as the share of all single familiy dwellings in the total dwelling stock (both heated and unheated dwellings). Single family dwellings include row house and townhouses, but not mobile homes. In the construction of these figures some problems exist with respect to comparisons between different data sources; for example uncertainty pertains to proper counting of farm houses, whether they are used for residential or commercial purposes.

Average dwelling area

The variable z_3 is defined as the weighted average dwelling area of total occupied dwellings in square meters.

Climate factor

The climate factor, z_5, is measured as the number of heating degree days using an 18 C base.

Share of central heating

The variable z_6 is defined as the share of all heated dwellings (included district heated dwellings) with a central heating system. Again, the exact definition of this variable can vary somewhat betweeen the various primary data sources. In general, central heating refers to a system which is able to heat the entire home, and in which the heat is distributed from a central point.

Interest rates

The data for the interst rates (ρ) are based on OECD's Main Economic Indicators, where we have used rates on government bonds for the different countries.

Equipment lifetimes

The variable d_h is defined as the expected lifetime of equipment of type h. At this point we have used US data, which were the only information available. Obviously, this is a weakness, since the duration of heating equipment may vary from country to country due to different qualities, efficiencies etc.

Gas network saturation

z_1 is defined as the share of dwellings in a specific country/year where gas is available as a heating fuel, in the sence that a household is located in a 'gas zone'. The definition is thus not restricted to dwellings actually hooked up to the gas grid. The data are collected from 'Le Marché Domestique du Gas', Données statistiques 1984.

A.3 Data limitations

Several problems were encountered when constructing a complete data set for the dynamic discrete-continuous choice model. One major difficulty was the lack of information regarding heating equipment prices. Other limitations of the data have directly influenced the actual model specification. One may for example refer to the fact that district heating is excluded as a separate technology in the model. At the same time, income per dwelling, the share of single family dwellings, the share of dwellings with central heating and average dwelling area are constructed using figures for total heated dwellings. The representation of 'central heating' as an independent variable is also a principal limitation of the model framework. A more satisfactory way of taking this variable into account would be to introduce explicitly a choice between a central and a non-central system in the model.

Instead of using average data for fuel shares in the various countries and years, *individual* observations of fuel choices would clearly have been preferable. The LBL data base does not provide sufficient information of *conversion* from on technology to another. Such data would obviously have been very useful for estimating the above kind of model.

References

Amemiya, T. (1981): Qualitative Response Models: A Survey. *Journal of Economic Literature*, Vol. 19.

Bohi, D.R. (1981): *Analysing Demand Behaviour. A Study of Energy Elasticities.* John Hopkins University Press.

Dagsvik, J.K. (1983): Discrete Dynamic Choice: An Extension of the Choice Models of Thurstone and Luce. *Journal of Mathematical Psychology*, Vol. 27.

Dagsvik, J.K. (1985): Markov Chains Generated by Maximizing of Components of Multidimensional Extremal Processes. Discussion Paper No. 12, Central Bureau of Statistics, Oslo.

Dagsvik, J.K. (1986): Stochastic Properties of the Market Wage and the Labour Supply Functions. A Theoretical Justification. Discussion Paper, Central Bureau of Statistics, Oslo (to appear).

Dagsvik, J.K., L. Lorentsen, Ø. Olsen and S. Strøm (1986): Residential Demand for Natural Gas: A Dynamic Discrete-Continuous Choice Approach. In *Golombek, R., M. Hoel and J. Vislie (eds.): Natural Gas Markets and Contracts.* North Holland Publishing Company, Amsterdam.

Dubin, J.A. and D. McFadden (1984): An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica*, Vol. 52.

Goett, A. and D.McFadden (1982): Residential End-Use Energy Planning System (REEPS). Report EA-2512, Electric Power Research Institute, California.

Hanemann, W.M. (1984): Discrete/Continuous Models of Consumer Demand. *Econometrica*, Vol. 52.

Heckman. J.J. (1981): Statistical Models for Discrete Panel Data. In *C.F. Manski and D. McFadden: Structural Analysis of Discrete Data with Econometric Applications.* MIT Press, Cambridge MA.

Hendry, D.F. (1974): Stochastic Specification in an Aggregate Demand Model of the United Kingdom.

Econometrica, May 1974.

MacRae, E.C. (1977): Estimation of Time-Varying Markov Processes with Aggregate Data. *Econometrica*, Vol. 45.

Maddala, G.S. (1983): *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge University Press, New York.

McFadden, D. (1973): Conditional Logit Analyses of Qualitative Choice Behaviour. In *P. Zarembka (ed.]: Frontiers in Econometrics.* Academic Press, New York.

Oren, S.S. and M.H. Rothkopf (1984): A Market Dynamics Model for New Industrial Products and its Application. *Marketing Science*, Vol. 3.

Ruderman, H. (1985): A Consumer Demand Model for Analysing the Effects of Efficiency Standards on

Appliance Choice and Residential Energy Consumption. Unpublished paper, Lawrence Berkeley Laboratory.

Schipper, L., Ketoff, A. and Kahane, A. (1985): Explaining Residential Energy Use by International Bottom - Up Comparisons. Unpublished paper, University of California, Berkeley.

Tiago de Oliveira, J. (1968): Extremal Processes: Definitions and Properties. Publ. Inst. Stat.

University of Paris, XVVII, 2.

Tiago de Oliveira, J. (1973): An Extremal Markovian Stationary Process. *Proceedings of the Fourth Conference in Probability Theory*, Acad. Romania, Brasov.

ISSUED IN THE SERIES DISCUSSION PAPER

No. 1 I. Aslaksen and O. Bjerkholt: Certainty Equivalence Procedures in the Macroeconomic Planning of an Oil Economy.

No. 3 E. Biørn: On the Prediction of Population Totals from Sample Surveys Based on Rotating Panels.

No. 4 P. Frenger: A Short Run Dynamic Equilibrium Model of the Norwegian Production Sectors.

No. 5 I. Aslaksen and O. Bjerkholt: Certainty Equivalence Procedures in Decision-Making under Uncertainty: an Empirical Application.

No. 6 E. Biørn: Depreciation Profiles and the User Cost of Capital.

No. 7 P. Frenger: A Directional Shadow Elasticity of Substitution.

No. 8 S. Longva, L. Lorentsen, and Ø. Olsen: The Multi-Sectoral Model MSG-4, Formal Structure and Empirical Characteristics.

No. 9 J. Fagerberg and G. Sollie: The Method of Constant Market Shares Revisited.

No.10 E. Biørn: Specification of Consumer Demand Models with Stochstic Elements in the Utility Function and the First Order Conditions.

No.11 E. Biørn, E. Holmøy, and Ø. Olsen: Gross and Net Capital, Productivity, and the Form of the Survival Function -

Some Norwegian Evidence.

No.12 J. K. Dagsvik: Markov Chains Generated by Maximizing Components of Multidimensional Extremal Processes.

No.13 E. Biørn, M. Jensen, and M. Reymert: KVARTS - A Quarterly Model of the Norwegian Economy.

No.14 R. Aaberge: On the problem of measuring inequality.

No.15 A-M. Jensen and T. Schweder: The Engine of Fertility - Influenced by Interbirth Employment.

No.16 E. Biørn: Energy Price Changes, and Induced Scrapping and Revaluation of Capital - A Putty-Clay Approach.

No.17 E. Biørn and P. Frenger: Expectations, Substitution, and Scrapping in a Putty-Clay Model.

No.18 R. Bergan, Å. Cappelen, S. Longva, and N. M. Stølen: MODAG A - A Medium Term Annual Macroeconomic Model of the Norwegian Economy.

No.19 E. Biørn and H. Olsen: A Generalized Single Equation Error Correction Model and its Application to Quarterly Data.

No.20 K. H. Alfsen, D. A. Hanson, and S. Glomsrød: Direct and Indirect Effects of reducing SO2 Emissions: Experimental Calculations of the MSG-4E Model.

No.21 J. K. Dagsvik: Econometric analysis of labor supply in a life cycle context with uncertainty.

No.22 Brekke K-A., E. Gjelsvik, B. H. Vatne: A dynamic supply side game applied to the European gas market.

No.23 Bartlett S. , J. K. Dagsvik, Ø. Olsen, S. Strøm: Fuel choice and the demand for natural gas in western european households.