Birds of a Feather Flock Together:

A Study of Doctor-Patient Matching

Geir Godager

Institute of Health Management and Health Economics, Health Economics Research Programme, University of Oslo

UNIVERSITY OF OSLO
HEALTH ECONOMICS
RESEARCH PROGRAMME

Working paper 2009: 3

HERO

Birds of a Feather Flock Together:

A Study of Doctor-Patient Matching

Geir Godager

27. April 2009

Health Economics Research Programme at the University of Oslo
HERO 2009

Keywords: GP services. Discrete choice. Willingness-to-pay. Health care demand.
JEL classification: C25, C83, D12, I11

Author:
Geir Godager, Institute of Health Management and Health Economics, Health Economics Research Programme, University of Oslo, P.O. Box 1089 Blindern, 0317 Oslo, Norway. Telephone: (+47) 22 84 50 29, fax: (+47) 22 84 50 91. E-mail: geir.godager@medisin.uio.no

Health Economics Research Programme at the University of Oslo
Financial support from The Research Council of Norway is acknowledged.
ISSN 1501-9071 (print version.), ISSN 1890-1735 (online), ISBN 978-82-7756-203-2

Abstract

In this paper we study individuals' choice of general practitioners (GPs) utilizing revealed preferences data from the introduction of a regular general practitioner scheme in Norway. Having information on relevant travel distances, we compute decision makers' travel costs associated with different modes of travel. Choice probabilities are estimated by means of nested logit regression on a representative sample of Oslo inhabitants. The results support the general hypothesis that patients prefer doctors who resemble themselves on observable characteristics: Individuals prefer GPs having the same gender and similar age.
Specialist status of GPs was found to have a smaller effect on choice probabilities than other attributes such as matching gender. When travel costs are calculated by means of taxi prices, the estimated willingness to pay for specialist status of a GP amounts to € 0.89 per consultation, whereas the estimated willingness to pay for having a GP with the same gender amounts to respectively € 1.71 and € 3.55 for female and male decision makers, respectively.

1. Introduction and background

When a patient consults a physician in an event of illness, the first of the physician's tasks is to reveal is the cause of illness, i.e. the diagnosis. A second task is to recommend an appropriate treatment and ensure that the patient is compliant with the treatment. Communication between the physician and the patient is an important element in both these processes. If the information transmission is efficient such that physician and patient are able to communicate easily and understand each other, the physician may be more likely to succeed in setting the correct diagnosis than if the converse was true. One may also argue that mutual confidence and unconstrained communication may cause the treatment to be more effective, as the degree of patient compliance is likely to be higher if the patient receives and understands the information relevant for the treatment. Often the doctor-patient relationship is described as a one of imperfect agency with the patient as the principal and the doctor as the agent. As described by Scott (2000, p.1179) the communicative ability of the matched doctor-patient unit is likely to affect the cost structure and the efficiency of a consultation, and transmission of information is thus likely to play a central role in meeting the objectives of the patient. The process of choosing a health care provider may thus be understood within the context of an agency paradigm, where part of the consumer's objectives is to affect the degree of imperfect agency, as suggested by Scott and Vick (1999). The consumer (principal) may mitigate agency imperfections by choosing a matching doctor (agent). We follow this idea and assume that patients prefer GPs who resemble themselves with respect to observable characteristics. This application of the old saying that "birds of a feather flock together" is shown to be a useful guide in the empirical specification where we model a representative decision maker's choice of GP within the random utility framework. The basic idea of this modeling framework is that a decision to choose a particular GP is considered the outcome of optimizing behavior, and a particular GP is chosen because the associated utility is higher than that of other alternatives.

The determinants of practice choice are examined in several studies, as reviewed in Scott (2000). Most earlier studies involving choice analysis and matching of GPs and patients consist of analysis of individuals' *stated* preferences with regard to hypothetical GPs. Examples of studies based on choice experiments are Scott and Vick (1998), Scott and Vick (1999), and Ryan et al. (1998). In these studies discrete choice experiments is applied to estimate the relative impact of different attributes of hypothetical GPs. While there are some obvious advantages with generating data in a controlled environment with appropriate sampling design, there are also drawbacks: The results are shown to be sensitive to the design and, in particular, the level of the attributes are shown to affect estimates of willingness to pay (Ryan and Wordsworth, 2000). Even though discrete choice experiments leave some important value judgments to the researcher, few studies use data on patients' *revealed*, rather than stated preferences. One example is Dixon et al. (1997), who examine the determinants of the rate at which patients left practices in three English health authorities. This study focuses on patients who revealed their preferences by switching practice without changing their home address. The main findings are that patients are more likely to leave a practice if it is small, if it is associated with longer travel distance and if it has shorter opening hours. They also find that 38 percent of the patients are registered with the practice closest to their home. Applying Norwegian data Lurås (2003) studies the consumers' ranking of GPs and find that individuals prefer GPs who are specialists as compared to GPs without specialist status. Other results are that consumers prefer a GP with the same gender, and that choice probabilities are found to be declining in the age difference between GP and patient.

An important feature of the latter studies is that endogenous sample selection is not accounted for even though one may argue that it is not obvious that individuals showing active switching behavior are representative for the general population. The present paper contributes to the literature by utilizing revealed preferences data in a setting where we are able to account for the potential selection bias resulting from endogenous sample selection. Having access to data describing the total population we are able to construct a representative sample of decision makers by means of the propensity score matching method. This material is well suited to study how attributes of GPs such as age, gender and specialization affect the individuals' choice of GP. The results from estimation of a nested logit model support the hypothesis that patients prefer GPs who resemble themselves with respect to observable characteristics. Individuals are more likely to choose a GP with the same gender, and the estimated choice probabilities are declining in the age difference.

The paper proceeds as follows. In Section 2 we give a brief description of the Norwegian reform of general practice. Data and sampling strategy is described in Section 3 while the econometric model is specified in Section 4. Results from estimation are given in Section 5 while Section 6 concludes and discusses the policy implications of the findings.

2. Institutional setting

The data used in this study is from Norway, a country with a national health service financed mainly through general taxation. A nationwide introduction of a regular general practitioner scheme in 2001 serves as a natural experiment providing detailed data on individuals' preferences for GPs. In order to implement this list patient system, every inhabitant was asked to return a response form ranking their three most preferred GPs in descending order. Since the submitted ranking information was intended to be used in the actual matching process forming each GP's patient list, this material constitutes a unique source of information on individuals' revealed preferences for GPs. Under the new scheme, more than 90% of the GPs are self-employed, with a payment system consisting of 30% per capita payment from the municipalities and 70% fee for service payment. The latter includes out of pocket payment from patients paying a fixed fee per consultation (€ 14.70 in 2001), with an annual ceiling. A special feature of Norwegian general practice is that two types of general practitioners exists: some have status as *specialist in general medicine*, the remaining do not have this status. GPs with specialist status are entitled to a higher consultation fee. The additional fee (€ 6.80) is financed by the National Insurance Administration. In order to achieve the formal specialist status the physicians are required to have more than four years of work experience in general practice, one year of experience from an inpatient or outpatient hospital department, and further, they need to fulfil a post-graduate education programme. This programme consists of courses, seminars and supervision from a senior GP. If one believes that more education adds to GP quality, specialist status may be considered to be an observable indicator of quality. Admittedly however, knowledge on specialist status of GPs is information that most likely is not acquired by every decision maker. In the same way that we expect costs to affect choices in situations were costs are hard to calculate[1], it is meaningful to investigate the impact of this attribute on choice probabilities. The reason is that the aim is to model the behavior of a representative decision maker. In summary, the market under consideration may thus be described as one where traded goods have observable quality differentiation and no consumer price variation, as the patients' out of pocket payments were the same for both types of GPs. An interesting question is then, does specialist status affect the demand for GP services, and if so, what is the magnitude of this effect?

3. Data and sampling strategy

Our data set is provided by the Norwegian Social Science Data Services. The observation unit is the individual inhabitant. All inhabitants in 14 Norwegian municipalities are included in the original data set. In this paper we will only use observations from inhabitants and GPs in the city of Oslo. The main reason for this decision is that an extract, containing the data from this densely populated metropolitan area gives more precise information on travel distances, compared to data from more rural areas where large geographical areas share the same postal code. As we know the residential addresses of consumers and practice addresses of GPs, a measure of the relevant travel distances in kilometers and travel time in hours can be added from a drive-time matrix.[2] One may argue that a limitation of this study is that we do not have exact information on the travel distances of each consumer. However, other methods of gathering information on travel distances would most likely also be imperfect. Further, the fact that Oslo has more than 400 unique postal codes, and that the distance matrix has recorded travel distances as short as hundred meters suggest that the measurement errors are small.

We are interested in studying the choice of sovereign consumers. Since parents are likely to choose the GP for their children we exclude observations of consumers

[1] Examples include phone rates, electricity tariffs and costs associated with car travel, etc.

[2] The private company *Infomap Norway* has collected actual travel distances and travel times associated with travel by means of a "light truck" on public roads between centers of the postal code areas.

younger than 18. After the exclusion of some observations where relevant information was missing, our sample has 401999 unique observations, of which 68% participated in the choice process. Descriptive statistics of the decision makers are given in Table 1. In the left column we give a description of the adult population of Oslo residents. The variable UNEMPLOYED is a dummy variable equal to 1 if an individual received any unemployment benefits in the period 2000-2002, and we se that see that 10% of the adult population has received such benefits. The variables NET WEALTH and INCOME consist of 10 groups categorized according to the deciles in the 14 municipalities. From the statistics on variable NET WEALTH we see that 10.5% of the population has a net wealth lower than the first decile, and we see that 10.3% has a net wealth between the first and second decile.

Only observations of individuals who returned the response forms, henceforth referred to as *participants*, can be used when estimating our choice model in Section 5. Individuals who did not take part in the GP choice process, henceforth referred to as *non-participants*, will therefore be excluded. As can be seen by comparing the three columns in Table 1, the consumers who participated in the choice process do not seem to be a representative sample of the inhabitants in Oslo.[3] We observe that a larger share of females returned their GP preferences as compared to males. We also observe that individuals with many years of schooling and high income are over-represented among participants, while younger individuals and people born in a foreign country, and people who have received unemployment benefits in the years 2000-2002 is clearly under-represented. The situation at hand has similarities with the sample selection situation described by van de Ven and van Pragg (1981). They study the demand for deductibles in private health insurance applying survey data where a large share of individuals returned incomplete questionnaires. They develop a two part binary probit model with endogenous sample selection in order to address the issue that the unobserved, and hence omitted, variable "expected medical expenses" is likely to relate both to the probability of completing the questionnaire, and to the probability of preferring a health insurance with a deductible. In the current situation one might suspect that the decision maker's state of health is related both to the probability of submitting provider preferences, and to the

[3]Confront Table A.1 in the appendix for a description of geographic representation in Oslo

Table 1: Descriptive statistics for exogenous variables
Population versus a self-selected and a corrected sample

	sample		
Variable	population N=401999 Proportion	self-selected n=15000 Proportion	corrected n=15000 Proportion
Female	0.522	0.581	0.520
Unemployed	0.101	0.082	0.104
Non-nordic	0.154	0.123	0.156
Schooling			
1-7 Years	0.006	0.005	0.006
8-10 "	0.138	0.148	0.145
11-12 "	0.220	0.237	0.228
13 "	0.206	0.199	0.207
14 "	0.023	0.023	0.022
15-17 "	0.251	0.252	0.241
18-19 "	0.098	0.099	0.096
20+ "	0.008	0.009	0.008
age			
30-40	0.238	0.207	0.238
40-50	0.171	0.177	0.170
50-60	0.147	0.174	0.149
60-70	0.084	0.106	0.084
70+	0.137	0.166	0.137
net wealth deciles†			
1	0.105	0.089	0.103
2	0.103	0.096	0.103
3	0.101	0.088	0.098
4	0.100	0.088	0.104
5	0.101	0.083	0.101
6	0.098	0.096	0.102
7	0.097	0.109	0.099
8	0.096	0.118	0.101
9	0.097	0.116	0.094
income deciles†			
1	0.100	0.081	0.096
2	0.094	0.090	0.098
3	0.097	0.095	0.094
4	0.097	0.105	0.102
5	0.098	0.104	0.098
6	0.099	0.104	0.100
7	0.101	0.109	0.101
8	0.102	0.106	0.103
9	0.104	0.106	0.103

† Deciles are calculated from the individual observations from 14 representative municipalities included in the original file. Decile1 refer to proportion of individuals with wealth/income less than Decile1. Decile2 refer to proportion of individuals with wealth/income between decile 1 and 2, etc.

relative valuation of the various attributes of GPs, such as GPs' specialist status. The empirical model is set up to model the decisions made by a "representative decision maker". If estimation is performed on a random sample from within the subset of self selected participants, the result may be biased coefficients or coefficients with an unclear interpretation. If the estimate of coefficients and the average willingness-to-pay is to have a meaningful interpretation, it is important that the decision makers included in the estimation sample really are representative for the population. As we are considering the choice between a large number of alternatives, the binary choice selection model considered by van de Ven and van Pragg do not seem applicable to the situation at hand. However, as we have a large number of observations and detailed information on the characteristics of both participants and non-participants we have the opportunity to *generate* a representative sample. Following Rosenbaum and Rubin (1983), we generate a representative sample of Oslo inhabitants by applying the method of propensity score matching, replacing non-participants with participants having approximately the same predicted participation probability.[4] The procedure may be described as follows: Let S denote the set of Oslo inhabitants, consisting of both participants and non-participants, expressed by $S = S_p \cup S_{np}$.

1. Estimate the probability of participation applying the total population, S, and calculate the predicted participation probability $\hat{\rho}_{is}$, $i = 1 \ldots 401999, s = p, np$
2. Draw a random sample $s \subset S$ of n individuals and obtain a sample of both participants s_p and non-participants s_{np}.
3. Replace the sampled non-participants, s_{np}, pairwise with participants who:
   (i) Are included in S_p but not included in s_p , and
   (ii) have approximately the same propensity score as the non-participants they are replacing: $\hat{\rho}_{inp} \approx \hat{\rho}_{jp}$

The results from the estimation of the participation probabilities are given in Table A.2 , and the details from the matching routine is described in Table A.3 in the appendix. By comparing the means in the third column of Table 1 with

[4] Representative samples can be achieved by beans of stratified sampling. Even though this is a simple approach with a small number of strata, it is not feasible in our situation where the aim is to account for a larger number of characteristics. The reason is that the number of distinct strata becomes unmanageable as the number of variables, or categories within each variable, increase: With 2 categories and V variables there are 2^V distinct strata.

the corresponding means of the population we see that a more balanced sample is achieved.

Table 2: Descriptive statistics for GPs. N=437

Variable	Mean	Std.dev
specialist	0.53	0.50
GP born in Norway	0.80	0.40
femaleGP	0.38	0.49
ageGP	47	7
marriedGP	0.66	0.47

The decision makers' choice menu consists of 437 alternative GPs meaning that 437 GPs have been ranked as the most preferred GP by at least one inhabitant.[5] In Table 2 we describe variables at the level of the GP. We observe that 53% of the GPs in Oslo are specialists in general medicine, and that 80% of the GPs in Oslo are born in Norway. Further, the average age of GPs in Oslo is 47 years and 38% of the GPs are females, and 66% of the GPs are married.

Since travel is costly, we expect that GPs with practices that are located close to the consumer's residential address are preferred to GPs located further away. We expect, ceteris paribus, the choice probabilities to be decreasing in travel time and travel distance. In order to achieve a monetary measure of the travel costs, a set of prices for distance and time is needed. A high-cost and a low-cost mode of travel is suggested, corresponding to travel by means of taxi and travel by means of private car. The fare schedule of the biggest taxi company in Oslo is used to get costs associated with taxi travel. To compute the costs associated with travel by means of private car a cost estimate of € 0.40 per kilometer is applied, which also corresponds to the reimbursement rate used by the Norwegian public sector to compensate employees for using their own car on official business.

The decision makers' own time is also part of travel cost. The "shadow price of time" is of course an individual specific variable and likely to be dependent of age, health and employment status. This information is not available at the level of the individual. A measure of the value of time spent on travel, as estimated

[5]We thus ignore the small number of GPs not ranked as number one by any of the decision makers.

by the Norwegian Institute of Transport Economics (Killi, 1999), is applied as the monetary cost of the decision makers' time use, although using such an aggregate is of course not beyond critique. The formulas for calculating travel costs are presented in Table 3. The travel costs associated with traveling to the GP is multiplied with a factor of 2, since patients travel both back and forth.[6]

Table 3: Formulas for calculating travel costs

Applying the fare schedule (TAXI), and a reimbursement schedule for the public sector (CAR).

Mode	Prices			formula
	start fee†	(€)/km	(€)/hour‡	
TAXI	(0) 4.10	1.30	6.80	travelcosts = 2×[4.10+1.30*km+ 6.80*hrs]
CAR	0	0.40	6.80	travelcosts = 2×[0.40*km + 6.80*hrs]

† Start fees are set to zero when distance is zero ‡ Inflation adjusted values for "time spent on travel" are from Killi (1999).

For given prices, the travel costs is a linear function of distance and time. Traveling to the GP by taxi is of course more expensive than traveling by own car, as the kilometer price is more three times as high. An equally important issue, however, is the fact that these two modes of transport have different cost *structures* as there is a starting fee required for each taxi trip.

In order to follow the idea that consumers prefer GPs who resemble themselves on observable characteristics, our representative utility function specified in the next section will include variables interacting characteristics of the alternative GPs with corresponding characteristics of decision makers. In Table 4 we describe the suggested interaction variables using the corrected sample of decision makers[7]. We compute the *absolute value* of the age difference between the patient and the GP, AGEDIFFERENCE. We see that the average age difference between consumer and the selected GP is 16 years. Since an increase in AGEDIFFERENCE implies that that the patient and GP are more different, we expect AGEDIFFERENCE to have a negative effect on choice probabilities. The dummy variable GENDER_{ff} (GENDER_{mm}) is equal to one when the female (male) decision maker and the GP have the same

[6]I am grateful to Sverre A. C. Kittelsen for pointing this out.

[7]Surprisingly, 7 individuals had selected a GP in one of the other 13 municipalities in the original data set. These individuals are excluded, and the corrected sample used for estimation in the following sections contains 14993 individuals

gender, and zero otherwise. We see that 23% of sample are men who chose a male GP while 38% are women who chose a female GP. In other words, 61% chose a GP with the same gender and 39% selected a GP with different gender. We expect GENDER$_{ii}$ to have a positive effect on choice probabilities. The mean travel distance between decision makers and the chosen GP is 2.64 kilometers and the mean travel time is 0.04 hours. We also see that the mean travel cost associated with travel by private car is € 1.42 whereas the mean travel cost associated with travel by taxi is € 7.40.

TABLE 4: DESCRIPTIVE STATISTICS FOR THE DECISION MAKER AND CHOSEN GP. INTERACTION VARIABLES. N=14993

VARIABLE	MEAN	STD.DEV	MIN	MAX
AGEDIFFERENCE	15.55	10.51	0	58
GENDER$_{ff}$	0.23	0.42	0	1
GENDER$_{mm}$	0.38	0.49	0	1
KILOMETERS	2.64	3.15	0	25.10
HOURS	0.04	0.05	0	0.37
COST CAR (€)	1.42	1.68	0	13.08
COST TAXI (€)	7.40	4.93	0	36.56

4. Random utility and the nested logit model

The choice of GP is a qualitative choice. Due to computational feasibility and convenience the most popular class of qualitative choice models is logit. The nested logit model to be derived here is a generalization of the multinomial logit (MNL) model described by McFadden (1974), and sometimes named McFaddens choice model. We denote by U_{nj} the utility consumer n obtains when selecting GP j.[8] Utility is equal to the sum of a component, V_{nj}, that is a function of variables that are observable and often called representative utility, and a component, ε_{nj}, that is unobservable and random, and we have:

$$U_{nj} = V_{nj} + \varepsilon_{nj} \tag{1}$$

[8]This deduction follows closely that of Train (2003, p. 81-85)

The crucial part of the assumptions underlying the standard MNL model is that the random factors, ε_{nj}, are uncorrelated over alternatives, as well as having constant variance across alternatives. In the context of this paper these assumptions would require unobservable factors related to alternative GPs located in the same neighborhood to be uncorrelated and have the same variance. One may argue that two GPs located in the same neighborhood are likely to be closer substitutes as compared to two GPs located in different areas. This kind of reasoning suggests that it may be appropriate to specify a nested logit model where GPs who are close substitutes are considered to belong to the same nest. Fortunately, it is straightforward to relax the restrictive assumptions underlying the standard MNL model and specify a nested logit model where the MNL model is included as a special case. The challenge is that an infinite number of nested logit models could be specified to represent the situation at hand, as any given city can be divided into geographical areas in an infinite number of ways. In particular, if one choose to specify a nest structure with a small number of large areas with many alternatives in each nest, one are less likely to place GPs that are close substitutes in different nests. On the other hand, one is more likely to place GPs that are *not* close substitutes in the same nest.

We now let the set of alternative GPs J be partitioned into K subsets denoted $B_1, B_2, ..., B_K$, $j \in B_k$; $k = 1, \ldots, K$ and refer to the K subsets as *nests.* The utility consumer n obtains when selecting GP j in nest B_k is equal to the sum of the deterministic and stochastic part of utility as expressed by (1). The nested logit model is obtained by assuming that the ε_{nj}'s has a joint cumulative distribution given by

$$\exp\left(-\sum_{k=1}^{K}\left(\sum_{j\in B_k} e^{-\varepsilon_{nj}/\lambda_k}\right)^{\lambda_k}\right)$$

The parameter λ_k indicates the degree of independence in unobserved utility between alternatives within nest k. A higher value of λ_k indicates greater independence and less correlation. When $\lambda_k = 1$ for all k, representing independence among all the alternatives in all nests, the nested logit model reduces to the MNL model. Testing the hypothesis $\lambda_k = 1$ for all k is thus a valid test for the appropriateness of the MNL model.

In this paper V_{nj} is specified as a linear function of observable variables:

$$V_{nj} = \boldsymbol{X}_{nj}\boldsymbol{\beta} + \boldsymbol{Z}_j\boldsymbol{\gamma}$$

where $\boldsymbol{X}_{nj}$ and $\boldsymbol{Z}_j$ are vectors of explanatory variables and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are vectors of the unknown parameters to be estimated. The latter parameters is assumed constant across nests, GPs and decision makers and may be interpreted as marginal utilities within the random utility framework. $\boldsymbol{X}_{nj}$ are explanatory variables interacting characteristics of the GP j with characteristics of consumer n, and $\boldsymbol{Z}_j$ are explanatory variables describing characteristics or attributes of GP j. In contrast with $\boldsymbol{X}_{nj}$, $\boldsymbol{Z}_j$ does not show any variation between decision makers, or in other words, there are no "within alternative" variation in $\boldsymbol{Z}_j$. These vectors will include the following variables:

$$\boldsymbol{X}'_{nj} = \begin{bmatrix} \text{GENDER}_{ff} \\ \text{GENDER}_{mm} \\ \text{AGEDIFFERENCE} \\ \text{TRAVEL COSTS} \end{bmatrix}, \; \boldsymbol{Z}'_j = \begin{bmatrix} \text{SPECIALIST} \\ \text{NORWEGIANGP} \\ \text{MARRIEDGP} \\ \text{AGEGP} \\ \text{AREA}_1 \\ \vdots \\ \text{AREA}_K \end{bmatrix}$$

We see that the $\boldsymbol{Z}_j$ vector include area indicators such that $z_{jk}^{nest} = 1$ if GP j is part of nest k. By estimating nest specific constants one ensure that the probability of choosing a GP within nest B_k is consistently estimated. Conditioned on $\boldsymbol{X}_{nj}, \boldsymbol{Z}_j$, the probability that consumer n choose GP i can be expressed as:

$$P_{ni} = P(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj};\, \forall i \neq j) = P(\varepsilon_{nj} < \varepsilon_{ni} + V_{ni} - V_{nj};\, \forall i \neq j)$$

A property of the nested logit model is that we get closed form expressions for P_{ni}. It can be shown that the probability of choosing GP i in nest B_k is given by:

$$P_{ni} = \frac{e^{V_{ni}/\lambda_k} \left(\sum_{j \in B_k} e^{V_{nj}/\lambda_k} \right)^{\lambda_k - 1}}{\sum_{l=1}^{K} \left(\sum_{j \in B_l} e^{V_{nj}/\lambda_l} \right)^{\lambda_l}}, \qquad i, j \in B_k;\; k, l = 1, \ldots, K$$

In the next section we present the results from a nested logit model where $K = 5$, that is, Oslo is divided in five nests by using postal codes. The Norwegian Mail Service refers to the two first digits in this code as the *postal code region*. In

Oslo there are 12 different postal code regions[9]: $01, 02, 03, \ldots, 12$. The postal code region are used to define the five areas referred to as WEST, NORTH, EAST_1, EAST_2, and SOUTH. The decision makers' choice set is the complete set of GPs that where actually available in Oslo when the regular GP scheme was implemented in June 2001, and all the decision makers are given identical choice sets. It should be noted that specifying a rank ordered logit model (Beggs, Cardell and Hausman, 1981) and utilizing the information on the alternatives ranked second and third was also considered. Although specifying such a model would allow us to extract more information from the data, extracting more information seems superfluous in the situation at hand.[10] Further, rank ordered logit models are vulnerable to heteroscedasticity (Hausman and Ruud, 1987) as choices of the alternatives with lower ranking are made more randomly. We therefore proceed and estimate a nested logit model by means of the maximum likelihood method available in the software STATA version 10.

5. Estimation and results

The results from nested logit regression are reported in Table 5. Most of the estimated coefficients are statistically significant. The results confirm the result from Lurås (2003) that GPs with specialist status have, cet. par., higher probabilities of being chosen than non-specialists. We also see that the estimated effect of the variables GP BORN IN NORWAY and MARRIEDGP are positive.
From the estimated effects of AGEGP (positive effect) and AGEDIFFERENCE (negative effect) we can make the interesting interpretation that consumers indeed do prefer GPs with similar age, but a GP who is older is preferred to a GP who is younger than oneself. An alternative interpretation of this result is that older GPs are preferred to younger GPs, and that the size of this positive effect of GP age is stronger the older the patient. The estimated effect of GENDER_{ff} and GENDER_{mm} has the expected sign, supporting the idea that consumers prefer GPs

[9] We ignore the postal code region 00 which is reserved for special addresses such as the royal castle.

[10] The author is also unaware of any standard software allowing for a nested specification of the rank ordered model.

Table 5: Results from nested logit estimation

No. of cases=14993. No. of Alternative GPs=437. Total No. of Obs. 6551941

	Taxi Travel Model		Car Travel Model	
Regressor	Estimate	Std.Err.	Estimate	Std.Err.
$\boldsymbol{X}_{nj}$ **variables**				
GENDER$_{ff}$	0.31	0.02**	0.27	0.02**
GENDER$_{mm}$	0.65	0.03**	0.54	0.02**
AGEDIFFERENCE	-0.02	0.00**	-0.02	0.00**
COSTS TAXI	-0.18	0.00**	-	
COSTS CAR	-		-0.51	0.01**
$\boldsymbol{Z}_j$ **variables**				
SPECIALIST	0.16	0.02**	0.15	0.01**
GP BORN IN NORWAY	0.18	0.02**	0.16	0.02**
MARRIEDGP	0.21	0.02**	0.18	0.01**
AGEGP	0.01	0.00**	0.01	0.00**
WEST (REF. CAT.)				
NORTH	-0.08	0.09	-0.19	0.09*
EAST1	-0.47	0.07**	-0.33	0.07**
EAST2	-0.72	0.09**	-0.70	0.08**
SOUTH	-0.85	0.09**	-0.56	0.08**
Dissimilarity parameters (λ_k)				
WEST	0.81	0.01**	0.72	0.01**
NORTH	0.86	0.03**	0.76	0.02**
EAST1	0.86	0.01**	0.73	0.01**
EAST2	0.78	0.02**	0.67	0.02**
SOUTH	0.88	0.02**	0.69	0.02**
Log likelihood	-68267.33		-68128.09	
P-value LR test for IIA	Prob > chi2 = 0.0000		Prob > chi2 = 0.0000	

(*) significantly $\neq 0$ at the 5 % level (two tailed test)
(**) significantly $\neq 0$ at the 1 % level (two tailed test)

with the same gender. It is also interesting to compare the *differences* in magnitude of male and female consumers' preferences for having a GP with the same gender as expressed by the difference in the estimated effect of the two variables GENDER$_{ff}$ and GENDER$_{mm}$. The results indicate that male consumers have stronger (p-value <0.01) preferences for having a GP with the same gender as compared to female consumers. We observe that the area dummies assigned to three of the city areas have significantly negative effect on choice probabilities. The interpretation is that a practice located in the reference category WEST is considered to be favorable by consumers. At the bottom of the table we observe the estimated values of the so called dissimilarity parameters referred to as λ_k. These parameters are in the range

[0,1] in nested logit models that are consistent with random utility theory[11]. The value of the dissimilarity parameters indicate the degree of intra nest correlation in unobserved utility, where values close to one imply low correlation and small values indicate high correlation. We see that the values of these parameters range from 0.78 (NORTH) to 0.88 (SOUTH), indicating that there are significant correlation in unobservable utility associated with alternatives within each nest. In the special case where $\lambda_k = 1$ for all k, the nested logit model collapse to the MNL model. At the very bottom of the table we see that the hypothesis that $\lambda_k = 1$ for all k is rejected, supporting the choice of a less restrictive nested logit model.

An application: Estimating the willingness to pay for GP attributes

TABLE 6: WTP ESTIMATES HIGH-COST AND LOW COST ALTERNATIVES

MODEL	HIGH-COST ALTERNATIVE: TAXI			LOW-COST ALTERNATIVE: PRIVATE CAR		
VARIABLE	WTP EST. IN €	95 % CONF. INT.		WTP EST. IN €	95 % CONF. INT.	
SPECIALIST	0.89	0.72	1.05	0.29	0.23	0.34
GP BORN IN NORWAY	0.99	0.79	1.20	0.31	0.25	0.37
MARRIEDGP	1.15	0.98	1.31	0.35	0.30	0.40
AGEGP	0.07	0.06	0.08	0.02	0.02	0.03
GENDER$_{ff}$	1.71	1.50	1.93	0.53	0.47	0.60
GENDER$_{mm}$	3.55	3.28	3.83	1.07	0.98	1.15
AGEDIFFERENCE	-0.10	-0.11	-0.09	-0.03	-0.03	-0.03

The vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ may be interpreted as marginal utilities. Having an estimate of the marginal utility associated with the travel costs, we may derive an estimate of the willingness to pay for attributes of GPs. This approach is often referred to as the *travel cost method*, a method more frequently used in environmental economics (Parsons, 2003). By definition, a decision maker's willingness to pay for an attribute such as specialist status is the increase in travel costs that keeps the decision maker's utility constant given that GP specialist status "change" from non-specialist to specialist. As described in Train (2003, p 43) we may take the total derivative of utility with respect to travel costs and specialist status and

[11]Dissimilarity parameters slightly larger than one are not necessarily inconsistent with random utility theory. Dissimilarity parameters may never be negative, however. For discussions of necessary and sufficient conditions for dissimilarity parameters to be consistent with random utility theory consult Börsch-Supan (1990) and Herriges and Kling (1996).

set this derivative to zero as utility is kept constant:
$\Delta\ U = \gamma\ \Delta$ SPECIALIST$+\beta\ \Delta$ TRAVEL COSTS$= 0$. Now we may solve for the change in travel costs that keeps utility constant for a change in specialist status:

$$\frac{\Delta \text{ TRAVEL COSTS}}{\Delta \text{ SPECIALIST}} = -\frac{\gamma}{\beta}. \tag{2}$$

We note that the willingness to pay is positive as the cost coefficient γ is negative. In Table 6 we have computed the willingness to pay estimates by using (2). The standard errors of these ratios, which are needed to calculate the confidence intervals, are obtained by means of the delta method (Wikipedia contributors, 2009). The disutility of the travel costs and the utility of the GP attributes is experienced at each consultation, and hence, the estimates presented in Table 6 denotes the willingness to pay *per consultation*. The willingness to pay estimates resulting from the Taxi Travel Model are higher than the estimates from the Car Travel Model. Still the willingness to pay for consulting a specialist in general practice appear to be low. When travel costs are calculated by means of taxi prices, the estimated willingness to pay for specialist status of a GP amounts to only € 0.89 per consultation, whereas the estimated willingness to pay for having a GP with the same gender amounts to respectively € 1.71 and € 3.55 for female and male decision makers, respectively.

6. Discussion and conclusion

The value or importance that decision makers attach to each attribute of the alternatives will in general vary. A limitation of the specified logit model is that it is unable to handle random taste variation. One might argue that some decision makers possess poor information on the concept of the specialist status, and hence that estimating the same $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ for all decision makers is a mis-specification. Although random taste variation can be incorporated in mixed logit models, estimation of such a model with the present choice set does not seem feasible due to the heavy computational burden. Estimation of a mixed logit model would most likely require a significant reduction in the number of alternatives. In this paper we have handled some elements of taste variation, by taking account of the possibility that GP attributes such as gender do not affect a male decision maker in the same way as a female decision maker, and by taking account of the fact that a young patient may value high physician age differently than an elderly patient.

Some of the consumers may have chosen a GP located close to their workplace, in order to combine everyday commuting with GP visits, and one may argue that *closeness to workplace* should be included as a GP attribute. Multi-purpose trips combining GP visits with commuting may imply that the computed travel costs are slightly exaggerated for some of these consumers. However, the presented model includes nest specific constant terms in the specification of representative utility implying that the choice probabilities, and the effects of travel costs, are identified by within-nest-variation in variables, and one may argue that unobservable effects such as "high density of work places" in certain areas are controlled for.

In order to assess the robustness of the estimated parameters and corresponding estimates of willingness to pay, several alternative models have been estimated. First, a model with 12 nests ($K = 12$) corresponding to the 12 postal code regions was estimated, and the results were compared with the above results. None of the estimated willingness to pay estimates were significantly different from the results presented here. The estimated dissimilarity parameters from the model with 12 nests were quite different, however: Several dissimilarity parameters were significantly larger than one, suggesting that the model might be inconsistent with random utility theory, hence the simple 5 nest model is presented in this paper. Second, the presented model was also estimated applying a random sample that was not corrected for endogenous sample selection. It is interesting to note that none of the estimated coefficients nor estimates of willingness to pay were statistically different. The implication of this result is that, even if there is evidence that the share of the population who took active part in the GP choice process differs from the share who remained passive, there is no evidence suggesting that their preferences for attributes of GPs are different.

There is evidence suggesting that consumers prefer GPs who resemble themselves on observable characteristics and it seems reasonable to conclude that the consumer's choice of GP is not random. Our estimates of the willingness to pay for consulting a specialist in general medicine seems to indicate that the willingness to pay is quite low and lower than the extra fee specialists in general medicine received at the time. An interpretation is thus that that the authorities' willingness to pay is higher than that of the patients. Several scenarios may lead to such an outcome. One particular scenario that is consistent with the presented results is one where the a higher consultation fee is motivated by specialists being closer substitutes to secondary care, and further, that specialists are expected to have lower referral rates compared to non-specialists. The specialist status is indeed valued by patients, and even more so by the authorities because fewer referrals to secondary care means lower health care costs. In other words we may not conclude that the situation at hand is one where the supply of specialist consultations are higher than what is socially optimal.

Since 2005, part of the extra fee specialists receive is paid by the patient, in the form of a € 3.30 patient co-payment. Since this co-payment rate is higher than the willingness to pay estimates presented here, an idea for future research would be to examine whether the introduction of a patient co-payment for consulting specialists in general medicine has affected the demand for the services of these specialists.

Acknowledgements

Thanks to Eline Aas, Erik Biørn, Anne Hvenegaard, Tor Iversen, Sverre A. C. Kittelsen, Hilde Lurås, Siri Fauli Munkerud and Erik Sørensen for suggestions that have contributed to improvements of the paper. Thanks also to participants at The 28th Meeting of the Nordic Health Economists' Study Group, Tartu, August 2007 and The 7th European Conference on Health Economics in Rome, July 2008, for helpful comments on a previous version of the paper. Some of the data applied in this project are provided by the National Insurance Administration and Statistics Norway. The data have been prepared for research purposes by The Norwegian Social Science Data Services. The author alone is responsible for all analyses and interpretations made in this text. Financial support from the Norwegian Ministry of Health and Care Services and the Research Council of Norway through the Health Economics Research Programme at the University of Oslo (HERO), is acknowledged. The author has no conflicts of interests with regard to this research.

References

Beggs, S., S. Cardell, and J. Hausman (1981): Assessing the potential demand for electric cars, *Journal of Econometrics* **16**:1-19

Börsch-Supan, A., 1990, On the compatibility of nested logit models with utility maximization, *Journal of Econometrics* **43**: 373-388.

Dixon, P., H. Gravelle, R. Carr-Hill and J. Posnett (1997): Patient movements and patient choice, Report for National Health Service Executive York Health Economics Consortium, York.

Hausman, J. and P. Ruud (1987): Specifying and testing econometric models for rank-ordered data, *Journal of Econometrics* **34**: 83-104.

Herriges, J. A. and C. L. Kling (1996): Testing the Consistency of Nested Logit Models with Utility Maximization, *Economics Letters* **50**: 3339.

Killi, M (1999): "Anbefalte tidsverdier i persontransport", (In Norwegian) TØI report **1999:459** Oslo.

Lurås, H., (2003): Individuals' preferences for GPs Choice analysis from the establishment of a list patient system in Norway. *HERO Working Paper* **2003: 5**.

McFadden, D., (1974): Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*,Academic press, New York.

Parsons, G.R. (2003): The travel cost model. In: Boyle, K., Peterson, G. (Eds.), *A Primer on Non-Market Valuation.* Kluwer Academic Publishers, Dordrecht.

Rosenbaum, P. R., and D. B. Rubin (1983): The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**: 41-55.

Ryan, M., E. Mclntosh and P. Shackley (1998): "Using conjoint analysis to assess consumer preferences in primary care: An application to the patient health card", *Health Expectations* **1**:117-129.

Ryan, M. and S. Wordsworth, (2000): "Sensitivity of willingness to pay estimates to the level of attributes in discrete choice experiments", *Scottish Journal of Political Economy* **47**: 504-524.

Scott, A. (2000): Economics of general practice. In A.J. Culyer and J.P. Newhouse (Eds.) *Handbook of Health Economics*, Volume 1, Elsevier Science, Amsterdam, 1175-1200.

Scott, A. and S. Vick (1998): Agency in health care. Examining patients' preferences for attributes of the doctor-patient relationship. *Journal of Health Economics* **17**: 511-644.

Scott, A. and S. Vick (1999): Patients, doctors and contracts: An application of principal-agent theory to the doctor-patient relationship. *Scottish Journal of Political Economy* **46**: 111-134.

Train, K. E., (2003): *Discrete Choice Methods with Simulation.* Cambridge university press, Cambridge.

van de Ven, W. and B. van Praag, (1981): "The Demand for Deductibles in Private Health Insurance: A Probit Model with Sample Selection", *Journal of Econometrics* **17**:229-252.

Wikipedia contributors, (2009): "Delta method", *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/w/index.php?title=Delta_method&oldid=270012701 (accessed March 11, 2009).

Appendix A:

Table A.1: Geographical representation of decision makers
Comparing the population means, with means from a random and a corrected sample

Postal Code Dummies	Population N=401999	random sample Participants only n=15000	corrected sample Participants only n=15000
	Mean	Mean	Mean
02**	0.081	0.077	0.076
03**	0.092	0.086	0.091
04**	0.101	0.098	0.101
05**	0.126	0.117	0.127
06**	0.161	0.171	0.168
07**	0.063	0.069	0.062
08**	0.040	0.042	0.036
09**	0.086	0.089	0.088
10**	0.055	0.049	0.052
11**	0.085	0.103	0.089
12**	0.054	0.056	0.055

Table A.2: Results from Logit estimation

Estimating the probability of participating in the GP choice process. No. of obs. = 401999

Regressor	Coeff. Est.	Std. Err.
Female	0.670	0.008 **
Education$_1$	0.093	0.063
Education$_2$	0.060	0.047
Education$_3$	0.265	0.047 **
Education$_4$	0.339	0.047 **
Education$_5$	0.340	0.052 **
Education$_6$	0.478	0.047 **
Education$_7$	0.518	0.048 **
Education$_8$	0.526	0.062 **
Missingedu	-0.554	0.049 **
Agecat$_2$	0.229	0.010 **
Agecat$_3$	0.614	0.012 **
Agecat$_4$	0.889	0.013 **
Agecat$_5$	1.224	0.017 **
Agecat$_6$	1.065	0.016 **
Net Wealth		
Decile$_1$	-0.373	0.017 **
Decile$_2$	-0.338	0.018 **
Decile$_3$	-0.376	0.018 **
Decile$_4$	-0.389	0.018 **
Decile$_5$	-0.427	0.019 **
Decile$_6$	-0.157	0.019 **
Decile$_7$	0.006	0.019 **
Decile$_8$	0.095	0.019 **
Decile$_9$	0.096	0.019 **
Total Income		
Decile$_1$	-0.173	0.018 **
Decile$_2$	-0.150	0.018 **
Decile$_3$	0.005	0.018 **
Decile$_4$	0.114	0.018 **
Decile$_5$	0.219	0.018 **
Decile$_6$	0.247	0.017 **
Decile$_7$	0.266	0.017 **
Decile$_8$	0.216	0.016 **
Decile$_9$	0.150	0.016 **
Unemployd	-0.376	0.012 **
cityarea$_2$	0.261	0.019 **
cityarea$_3$	0.221	0.018 **
cityarea$_4$	0.279	0.018 **
cityarea$_5$	0.206	0.017 **
cityarea$_6$	0.464	0.017 **
cityarea$_7$	0.464	0.021 **
cityarea$_8$	0.591	0.024 **
cityarea$_9$	0.498	0.019 **
cityarea$_{10}$	0.548	0.021 **
cityarea$_{11}$	0.730	0.020 **
cityarea$_{12}$	0.711	0.021 **
europe	-0.155	0.014 **
usacanada	-0.175	0.043 **
africa	-0.390	0.023 **
asia	-0.108	0.015 **
oceania	0.022	0.144 **
southamerica	-0.581	0.041 **
constant	-0.556	0.053 **
Log likelihood	-227485.91	
Pseudo R^2	0.0975	

Table A.3: Description of propensity score matching routine

Columns 1 and 4 records the estimated propensity score among non-participants in the random sample s with 15000 observations. Columns 2 and 5 reports the number of nonparticipants needed to be replaced. Columns 3 and 6 reports the number of matching candidates. We define a match when $|\hat{\rho}_{np} - \hat{\rho}_p| < 0.01$

$\hat{\rho}_{np}$	# to replace	# matching candidates	$\hat{\rho}_{np}$	# to replace	# matching candidates
.07	**1**	**0***	.52	118	2435
.10	1	10	.53	135	2475
.12	2	9	.54	99	2275
.13	1	20	.55	103	2319
.14	8	29	.56	134	2917
.15	3	30	.57	125	3214
.16	9	27	.58	117	3049
.17	7	63	.59	110	3322
.18	4	43	.60	128	3180
.19	6	55	.61	130	3293
.20	7	87	.62	114	4026
.21	12	95	.63	129	3873
.22	9	89	.64	107	3657
.23	20	118	.65	124	3915
.24	17	171	.66	124	4617
.25	17	127	.67	137	5144
.26	14	137	.68	137	5369
.27	29	215	.69	120	4678
.28	35	244	.70	100	4707
.29	28	236	.71	109	5906
.30	30	193	.72	122	6127
.31	34	282	.73	98	5654
.32	27	318	.74	103	5917
.33	30	431	.75	99	5810
.34	42	476	.76	114	7191
.35	54	511	.77	99	6383
.36	40	631	.78	87	6864
.37	51	621	.79	87	7095
.38	53	644	.80	128	8386
.39	48	744	.81	85	7354
.40	55	878	.82	72	7439
.41	61	995	.83	101	8715
.42	68	1159	.84	81	8059
.43	63	1136	.85	68	8034
.44	78	1125	.86	76	8579
.45	71	1226	.87	66	7627
.46	93	1602	.88	66	7097
.47	74	1577	.89	62	6696
.48	88	1762	.90	31	5084
.49	81	1753	.91	16	3867
.50	85	1955	.92	21	2718
.51	121	2376	.93	11	1247

* No match was found for the non-participant with propensity score 0.07. This particular observation was matched with a candidate with a propensity score of 0.08.