

Four Empirical Essays on the Market for General Practitioners' Services

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: 7

HERO

Four Empirical Essays on the Market for General Practitioners' Services

Geir Godager

Dissertation for the ph.d. degree Department of Economics University of Oslo

September 2009

Health Economics Research Programme at the University of Oslo HERO 2009

Keywords: GP services; discrete choice; willingness-to-pay; health care demand; health

care supply; general practice; patient shortage; dual job holding

JEL classification: C23; C24, C25, C33, C83, D12, D83, H51, H75, I11, I18, J22

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

Contents

	0.1	Preface A note on pile driving and doubt	4
	0.2	List of original essays	5
1	Intr	roduction	6
	1.1	Motivation	6
	1.2	Health economics and general practice	7
	1.3	Summary of essays	12
	1.4	Policy implications	17
	1.5	Limitations and suggestions for further research	19
	1.6	References	20
2	Essa	ay 1: Birds of a Feather Flock Together:	
	A S	tudy of Doctor-Patient Matching	22
3	Essa	ay 2: Does Quality Influence Choice of General Practitioner?	,
	An	Analysis of Matched Doctor-Patient Panel Data	46
4	Essa	ay 3: Dual Job Holding General Practitioners:	
	The	Effect of Patient Shortage	78
5	Essa	ay 4: Service Motives and Profit Incentives Among	
	Phy	rsicians	103

Acknowledgements

This work has been financed by the Norwegian Ministry of Health and Care Services and the Norwegian Research Council through the Health Economics Research Programme at the University of Oslo (HERO). I am sincerely grateful for the financial support. Thanks to the Norwegian Social Science Data Services (NSD) which has provided the data applied in these essays. NSD is responsible for neither data analyses nor interpretations. I am grateful to the Institute of Health Management and Health Economics, University of Oslo, for providing me with excellent working conditions.

I am indebted to a great number of people who have generously offered inspiration, encouragement, support and advice. Special thanks to Erik Biørn, my supervisor and co-author, to Tor Iversen, my secondary supervisor and co-author and to my co-authors Hilde Lurås and Ching-to Albert Ma. These individuals have made this dissertation possible. Erik and Tor encouraged and inspired me in every meeting. Hilde and Tor encouraged me to apply for doctoral studies. A great number of people have provided advice and suggestions for improvements. For this I would like to thank Eline Aas, Arna Desser, Sverre Grepperud, Anne Hvenegaard, Sverre A. C. Kittelsen, Ismo Linnosmaa, Karin Monstad, Siri Fauli Munkerud, Trond Olsen, Terje Skjerpen, Erik Magnus Sæther, Erik Sørensen and Gilberto Turati. Thanks also to my supportive family, friends and colleagues.

I dedicate this thesis to my wife, Anita, and children, $J,\,J$, J and $_J$.

0.1 Preface. - A note on pile driving and doubt

It is often argued that Karl Popper's philosophy of science is the basis for much of the research in the field of economics (Smith, 1998). With Popper's philosophy naive induction must give way to deductive logic. In addition, Popper has a modest level of ambition with regard to the goals of science: The scientific objective is "Truth", but we will never know whether we have found "the Truth", and it is not certain that science will ever achieve "the Truth". Our knowledge does not have solid foundation. In the words of Popper (1959, p. 111):

"The empirical basis of objective science has nothing "absolute" about it. Science does not rest upon solid bedrock. The bold structure of its theories rises, as it were, above a swamp. It is like a building erected on piles. The piles are driven down from above into the swamp, but not down to any natural or "given" base; and if we stop driving the piles deeper, it is not because we have reached firm ground. We simply stop when we are satisfied that the piles are firm enough to carry the structure, at least for the time being."

The aim of this thesis is to contribute to the accumulation of knowledge and understanding of mechanisms at work in the market for general practitioners' services by confronting theory with data. In Poppers terminology, this dissertation is hopefully contributing to some "pile driving" within the economics of general practice. There will always be room for doubt however, and importantly, anyone convinced that "the Truth" is revealed will hopefully experience some scepticism. In the words of Fuchs (1986) in a reply to a previous discussion on the phenomenon of supplier induced demand:

"...This ... has always reminded me of the story of the Frenchman who suspected that his wife was unfaithful. When he told his friend that the uncertainty was ruining his life, the friend suggested hiring a private detective to resolve the matter once and for all. He did so, and a few days later the detective came and gave his report: 'One evening when you were out of town I saw your wife get dressed in a slinky black dress, put on perfume, and go down to the local bar. She had several drinks with the piano player and when the bar was closed they came back to your house. They sat in the living

room, had a few more drinks, danced, and kissed.' The Frenchman listened intently as the detective went on: 'Then they went upstairs to the bedroom, they playfully undressed one another, and got into bed. Then they put out the light and I could see no more.' The Frenchman sighed 'Always that doubt, always that doubt.'"

0.2 List of original essays

- Essay 1: "Birds of a Feather Flock Together: A Study of Doctor-Patient Matching" (Godager, G.). Essay 1 is submitted to *Journal of Public Economics*
- Essay 2: "Does Quality Influence Choice of General Practitioner? An Analysis of Matched Doctor-Patient Panel Data" (Godager, G. and E. Biørn). Essay 2 is submitted to *Economic Modelling*.
- Essay 3: "Dual Job Holding General Practitioners: The Effect of Patient Shortage" (Godager, G. and H. Lurås). Essay 3 is accepted for publication in *Health Economics*.
- Essay 4: "Service Motives and Profit Incentives Among Physicians" (Godager, G., C.A. Ma and T. Iversen). Essay 4 is published in *International Journal of Health Care Finance and Economics*, **9**:1, 39-57.

1 Introduction

1.1 Motivation

This dissertation consists of empirical essays within the subject of health economics. There are four essays in applied micro-econometrics and, as data in Essays 2 and 4 have a panel format, econometric methods for panel data are applied. Tobit-type models for limited dependent variables are applied in Essays 3 and 4, Nested logit models for choice of general practitioner (GP) are applied in Essay 1 and in Essay 2 modeling and estimation procedures involving latent structural variables are applied. The market for GPs' services is the recurrent theme, and a common feature of the empirical modeling and estimation conducted in the essays of this dissertation is that latent variables play an important role.

There are several motivations for studying the market for GPs. GPs play a key role and constitute the cornerstone of the health care sector in Norway and other countries with a national health service. The GP is often a patient's first encounter with the health care sector. In Norway the GP also acts as a "gatekeeper", and a referral from the GP is necessary in order to receive specialized care. Further, decisions made by GPs have a large impact on public spending such as sick-leave benefits and drug reimbursements. Due to this pivotal role of general practice, any research providing policy guidance for the sector will potentially have noticeable welfare effects.

Studying this market may also provide advances in terms of enhanced understanding of economic behavior in general, and within the discipline of health economics in particular. Arrow's (1963) article describing various imperfections in the market for medical care is often considered to mark the founding of health economics (Culyer and Newhouse, 2000). Many of the peculiarities Arrow describes are, indeed, present in the market for GPs' services. Asymmetric information limits consumer sovereignty and creates challenges for designing appropriate contracts and payment mechanisms. Quality of services is difficult to observe and quantify, and optimal consumption of health care services is difficult to achieve. While the general research questions in health economics are relevant in the specific context of general practice, the conditions for knowledge accumulation seem favorable as well. The availability of detailed and disaggregated panel data enables identifi-

cation and quantification of the mechanisms in focus by applying a large variety of modeling and estimation methods. In addition, one may argue that economic theory is well suited to model individual behavior and that assumptions of rational decision makers are more realistic when describing individuals such as GPs than when describing institutions such as hospitals. Fuchs (2000) describes two related missions of health economists: providing valuable input into health policy and enhancing understanding of economic behavior. Both missions motivate research on the market for GPs.

Essays 1 and 2 in this dissertation focus mainly on factors influencing the patients' or consumers' decision to seek the services of a particular GP, while Essays 3 and 4 focus on factors influencing the GP's supply decision. The rest of the text proceeds as follows: The next subsection provides a brief introduction to the topic of this dissertation. Summaries of the four essays are given in subsection 1.3. Policy implications of the findings are presented in subsection 1.4. Limitations and ideas for future research are discussed in subsection 1.5. Complete versions of Essays 1-4 follow in sections 2-5.

1.2 Health economics and general practice

The health economics literature on general practice is growing. A detailed review is beyond the scope of this text but Scott (2000) and Iversen and Lurås (2006) offer extensive reviews of the economic research on general practice. In this literature the doctor-patient relationship is often described as one of imperfect agency with the patient as the principal and the doctor as the agent, and where important elements such as "Quality of Care" are noncontractible (McGuire, 2000), meaning that these elements can not form the basis for payment. In the context of general practice these noncontractibles are most likely of fundamental importance since diagnostic precision and treatment outcomes are determined, at least partly, by provider effort. Therefore regulators are concerned about noncontractibles when designing the payment system. The most prominent payment mechanisms in general practice are fixed salary, fee-for-service payment and capitation (with or without risk adjustment), and combinations of these. Results from the literature on GP behavior support the idea that the way in which services are paid for influences how services are delivered. As described in Iversen and Lurås (2006), salaried GPs receive the

same income irrespective of the number of treated patients, and this may result in low productivity. Further, while salaried GPs do not have incentives to over-supply certain tests and examinations, they do not have incentives to economize on the use of factor inputs. Fee-for-service payment seems to encourage the provision of services, and the resulting service volume can be higher than optimal. Under capitation payment the GPs internalize all costs associated with the treatment and, in contrast to fee-for-service payment, one is more likely to achieve cost-minimizing combinations of factor inputs. On the other hand, capitation payment may create incentives to minimize effort by engaging in unnecessary referring and prescribing in order to minimize own workload. In the UK, the introduction of the fundholding scheme in general practice encouraged GPs to internalize the costs of referring patients for specialist care. Under the fundholding scheme GPs were provided with a budget to cover drug costs and costs for elective surgery for their patients. Any budget surplus was retained in the practice; fundholding can therefore be regarded as an extended form of capitation. Dusheiko et al. (2006) found that the scheme contributed to a "downward pressure" on the number of secondary care admissions.

Newhouse (1996) refers to pure capitation or pure fee-for-service systems as "corner solutions" and suggests that the optimal payment system for health care providers, is a mixed system:

If pure fee-for-service results in overservicing and pure capitation in underservicing, the optimal scheme is a mixture.

The advantages of mixed payment mechanisms are an important result from the literature on industrial organization of health care markets; the current payment system for GPs in the UK may be an example of how results from research are put into practice: Currently, most of the GPs in UK are paid under a contract consisting of a mix of capitation, lump sum allowances, fee-for-service and payment according to the so called Quality and Outcomes Framework (Gravelle et al., 2008). The introduction in the UK of the Quality and Outcomes framework, which monitors outcome based on 146 quality indicators, signals an increasing focus on quality in the context of GPs' services.

Concerns for quality of services seem to be increasing in Norway as well. The reform in general practice, implemented in 2001, reflected the view that giving

GPs responsibility for a particular list of individuals, while simultaneously highlighting the patient's right to choose and switch general practitioners, would result in increased competition among GPs and improvements in accessibility and quality.

Studies based on data from the Norwegian Survey of Living Conditions conducted by Statistics Norway suggest that an improvement in accessibility has occurred after 2001: The average time a patient has to wait from booking a consultation until the consultation occurs has been reduced. The proportion of patients who receive an appointment within the same day doubled (from 11% to 22%) and the median wait was reduced from 4 to 2 days from 2001 to 2005 (Godager et al., 2007). Lurås (2007) describes the degree to which Norwegian patients express satisfaction with their GP along five quality dimensions, and find that most respondents are largely satisfied with the quality of services provided by their regular GP.

Latent variables

A characteristic of health economics is that variables of crucial importance are often unobservable to agents in the market and researchers studying the market; application of models involving latent variables is therefore prevalent (Jones, 2000, p. 268). Latent variables are also a common feature of the four essays in this dissertation. The motivation behind the latent variable specification is different across essays. In Essay 1 the latent variable approach is used to motivate the logit model and the random utility framework. The utility consumer n obtains when selecting GP j from within set, J, is of course not observable, what is observed is only the choice that is made. Still it is reasonable to infer that the latent utility of the consumer associated with the other alternatives is lower.

The motivation behind the latent variable specification in Essay 2 is somewhat different. For some variables it is not obvious what scale of measurement is relevant; the quality of health care services is an example of this type of variable. In addition to lacking a scale of measurement, quality is often considered to be unobservable to agents in the market as well as researchers. In the literature, quality of services is frequently represented or measured by applying *proxies* or *indicators* of quality such as the in-hospital mortality rate. There is a growing literature studying the effect of various market characteristics on observed mortality rates (Gaynor, 2006). The rationale for applying mortality rates as a measure of quality is often not

discussed in these papers. One may argue however that the mortality rate is an outcome measure, that is, the result of a process where quality is actually one of several inputs. One may argue that quality of health care services is related not only to the length of life but also to the quality of life, and hence that mortality rate is only one of several possible indicators of quality. Further it is obvious that two health care providers with the same level of observed mortality rates are not necessarily of equal quality, as both case mix and the rate of hospital infections may, of course, differ. When several indicators of quality is available, the question arising is, how should one weigh these indicators in order to interpret variations in unobservable quality? This is one of the issues addressed in Essay 2 in which the relationship between health care quality and demand is modeled by applying LISREL modeling and estimation procedures. (Jöreskog, 1977, Aigner et al., 1984)

The censored regression models applied in Essays 3 and 4 are also based on latent variable specifications. The censored regression model proposed by Tobin (1958) was developed to model the demand for durable goods, but these so called tobit models also proved useful in the context of models for labor supply (Moffitt, 1999). The special feature of tobit type censored regression models is that the latent variable is observable when the realized values are in a certain range or exceed a threshold. Labor supply is inherently non-negative, and the latent labor supply is unobservable when decision makers choose not to participate in the labor market. In Essay 3 and 4 institutional details from prevailing regulations are of fundamental importance in specifying empirical models. According to prevailing regulations, the municipalities can require GPs to work up to 7.5 hours per week with community health service tasks. In standard labor markets labor supply is censored at zero, in this case we must allow the censoring threshold to be GP specific and allow it to vary in the interval [0, 7.5] hours. Essay 4 differs from the third essay in that a censored regression model is estimated by applying panel data, and a latent variable specification is used to represent unobserved heterogeneity of physicians in this random effects tobit model. In summary, all the 4 essays in this dissertation constitute examples of what can be achieved if one is willing to make assumptions about the structure of what we do not observe.

Institutional setting and data sources

Norwegian data are applied in the four essays in this dissertation. Norway has a national health service, mainly financed by general taxation. In all four essays some of the data are from the Norwegian primary physician database, which describes characteristics of each GP and each GP's patient list. This database was established in connection with the introduction of the list system in general practice in 2001. Under the new system GPs are paid a fee-for-service reimbursement and a capitation per listed patient. Besides providing primary care, GPs act as gatekeepers, and a referral is required for consultations with health care specialists. Patients pay a copayment of about \in 15 per consultation with their GP. National insurance covers copayments if expenditures for physician services and medicines within a year exceed a ceiling.

The Norwegian primary physician database is administered by the Norwegian Social Science Data Services and provides information about individual GPs. For Oslo, the capital of Norway, and 13 other municipalities the database also includes more detailed information at the level of the individual inhabitant. For these 14 municipalities the registry data include information on address, wealth, income and education in addition to the individual's revealed preferences for GPs. This information is available due to the significant effort made by the Norwegian authorities to involve individuals in the introduction of the list system: In order to implement the reform and list each inhabitant with a GP, 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 for use in the actual matching process forming each GP's patient list, this material constitutes a unique source of information about individuals' revealed preferences for GPs. This ranking information is applied in different ways in Essay 1 and Essay 2. Survey data describing the GPs' participation in the community health service was collected in two waves, in 2002 and 2004. The survey data were merged with registry data from the Norwegian primary physician database. Survey data are applied in Essays 3 and 4.

1.3 Summary of essays

Summary of Essay 1: Birds of a Feather Flock Together: A Study of Doctor-Patient Matching

In this paper we study individuals' choice of GPs utilizing revealed preferences data collected during the introduction of the regular general practitioner scheme in Norway. The individual consumer's choice of GP is modeled within the random utility framework, and the main hypothesis is that patients prefer doctors who resemble themselves on observable characteristics. Using information about relevant travel distances, we compute decision makers' travel costs associated with two different modes of travel: travel by taxi and travel by means of private car. Choice probabilities are then estimated by means of nested logit regression. The nested specification is chosen because GPs located in the same neighborhood are expected to be closer substitutes compared to GPs located in different neighborhoods.

An important feature of earlier studies of GP choice applying revealed preference data is that the possibility of endogenous sample selection is not taken into account. One may argue that this is a limitation because it is not obvious that individuals showing active interest in engaging in provider choice are representative of the general population. In the present paper we observe that a larger share of females than males returned their GP preferences. We also observe that individuals with many years of schooling and high income are over-represented among individuals submitting provider preferences, while younger individuals, people born in a foreign country, and people with recent unemployment spells are clearly underrepresented. The present paper contributes to the literature by utilizing revealed preference 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. Hence, the nested logit regression is estimated applying 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 GP's gender and country of birth: When travel costs are calculated by means of taxi prices, the estimated willingness to pay for consulting a GP with the same gender is \in 1.71 and \in 3.55 for female and male decision makers, respectively, whereas the estimated willingness to pay for specialist status of a GP is \in 0.89 per consultation. The corresponding willingness to pay estimates are smaller when travel costs associated with travel by private car are applied.

At the time our data were collected, GPs who were specialists in general medicine received an additional fee of \leq 6.80 per consultation from the National Insurance Scheme, while the patients' out of pocket fee for consulting a specialist in general medicine was the same as for consulting a non-specialist. The results thus indicate that average willingness to pay for consulting a specialist is considerably lower than the additional fee that specialists receive from the National Insurance Scheme.

Summary of Essay 2: Does Quality Influence Choice of General Practitioner? An Analysis of Matched Doctor-Patient Panel Data

The aim of this paper is to investigate empirically whether the demand facing a general practitioner (GP) responds to the quality of the provided services. The impact of quality on the demand for health care providers has important implications for the industrial organization of health care markets. In this paper we study the consumers' choice of GP assuming that they are unable to observe the true quality of GP services. A panel data set for 484 Norwegian GPs, with summary information on their patient stocks, provides the opportunity to identify and measure the impact of GP quality on demand, accounting for patient health heterogeneity in several ways. Two kinds of models are considered: a panel data model with latent heterogeneity related to GP quality and a multi-equation LISREL type of model, including both GP quality and the health of the stock of persons on the GP's list as latent variables, both of which are assumed to affect demand as well as other observed variables. The patient excess mortality rate at the GP level is one indicator of quality.

An important issue addressed in this paper is the potential selection occurring as a result of heterogeneity of patients' skill or ability to collect and process available quality information. This heterogeneity may induce selection mechanisms

resulting in systematic differences in patient morbidity and patient mortality between GPs with different levels of quality. We have no prior knowledge of the direction in which this selection mechanism might operate. On the one hand, less healthy consumers, with a high expected mortality rate, might be particularly concerned about their choice of GP and as a result more willing to collect information than the average consumer. This may contribute to increasing average mortality rate among the patients listed with high-quality GPs. On the other hand, consumers who are healthier, more resourceful and have low expected mortality may be particularly able to collect and process such information. This could reverse the selection mechanism, i.e., lowering the average mortality rates of the persons listed with high-quality GPs. Consequently, a priori, we cannot postulate that patient heterogeneity will result in a difference in the average health status of patients of high-quality vs. low-quality GPs or the direction of the difference, if it exists. However, failing to control for differing aggregate health status of listed patients may result in a simultaneity bias and/or an excluded variable bias when estimating the effect of quality on demand. Our LISREL approach separates the effect of quality on excess mortalities from the effect of health at the level of the GP through the exclusion restrictions imposed on the measurement equations for Quality and Health: six variables describing the GP are included in the measurement equations for Quality, but excluded from the measurement equations for Health, while three variables are included in the measurement equations for Health but excluded from the measurement equations for Quality. We estimate the effect of the quality variable on the demand for each GP's services. Our results, obtained from two different econometric model versions, indicate that GP quality has a clear positive effect on demand.

This paper adds to the literature in several ways. First, no previous empirical studies seem to have considered the demand effects of quality in the market for GPs. Second, in the current literature the relationship between demand and various indicators of quality, such as mortality rates, failure rates or hospital type, and other independent variables, are estimated separately. The paper contributes to the literature by simultaneously estimating the relationship between demand and quality and the way in which quality becomes manifest via indicators, applying linear structural equation modeling (LISREL) and estimation methods. Taking

this approach, we acknowledge both the multidimensional nature of the quality concept, and that it may be more appropriate to interpret outcome measures such as mortality rates or failure rates as indicators of quality, rather than as direct measures of quality. Third, our econometric model provides a method to separate the effect of quality on outcome measures from the effect of patient health.

Summary of Essay 3: Dual Job Holding General Practitioners: The Effect of Patient Shortage

In 2001, a list patient system with capitation payment was introduced in Norwegian general practice. After the allocation process by which each inhabitant was listed with a GP, a considerable share of GPs had fewer persons listed than they would have preferred. In this paper we examine whether GPs who experience a shortage of patients, to a larger extent than other GPs, seek to hold a second job in the community health service even though the wage rate is low compared to the payment in general practice. While privately practicing GPs are responsible for providing general medical services to persons listed in their practices, the municipalities' community health service consists of certain segments of the primary-care sector that are not part of the domain of general practice:

- Administration of the primary health-care sector;
- Public health services such as the preparation of infectious disease plans;
- Certain medical tasks such as routine examinations of infants and children at childcare centers and schools, and regular medical care at nursing homes, prisons and other institutions.

These tasks in the community health service are carried out by GPs having parttime positions in the municipality.

Assuming utility maximization, we model the effect of patient shortage on a GP's decision to contract for a second job in the community health service. The model predicts a positive relationship between patient shortage and participation in the community health service.

Applying data from 387 GPs practicing in 26 municipalities in Norway and two districts in Oslo (the capital city), we estimate a censored regression model, taking account of labor supply as a censored variable. In standard labor markets, individuals who choose not to offer their working capacity to the market will be observed with a labor supply of 0 hours, i.e. the variable is censored at zero. Because the municipalities can require GPs to contract for up to 7.5 hours per week we must allow the GP's specific censoring threshold to vary in the interval [0, 7.5] hours.

The results indicate that GPs experiencing a shortage of patients contract for more hours with the community health service, and that the longer the constrained GPs' lists are, the fewer hours are contracted for. The estimated marginal effect of patient shortage is 1.72 hours per week. The effect of wage rate is statistically significant and the coefficient has the expected sign. The estimated marginal effect of wage rate indicates that increasing the wage rate by one € increases the GPs' observed labor supply by 0.154 hours per week. This result corresponds to conditional supply elasticity with respect to wage rate of 0.32. When standard errors are calculated by means of the delta method, the 95% confidence interval of the conditional supply elasticity is [0.018,0.629], which is comparable in magnitude to physician labor supply elasticities found in earlier studies. We also observe a significant effect of being a specialist in community medicine, and as expected, specialists in community medicine supply more hours than GPs without this speciality.

Summary of Essay 4: Service Motives and Profit Incentives Among Physicians

Essays 3 and 4 are thematically related. The number of hours supplied by the individual GP to the community health service is studied in both essays. The two essays approach the GPs' supply of hours to the community health service from two quite different angles. In this essay the point of departure is that GPs who work voluntarily in the community health service for low pay may derive utility from performing these tasks. We model physicians as health care professionals who care about their services and monetary rewards. These preferences are heterogeneous. Different physicians trade off the monetary and service motives differently, and therefore respond differently to incentive schemes. Our model reflects the Norwegian health care system. First, each private practice physician has a patient list, which may have more or fewer patients than he desires. The physician is

paid a fee-for-service reimbursement and a capitation per listed patient. Second, a municipality may obligate the physician to perform 7.5 hours per week of community services. Our data consist of an unbalanced panel of 435 physicians, with 412 physicians for the year 2002, and 400 for 2004. A physician's amount of gross wealth and gross debt in previous periods are used as proxy for preferences for community service. First, for the current period, accumulated wealth and debt are predetermined. Second, wealth and debt capture lifestyle preferences because they correlate with the planned future income and spending.

The main results show that both gross debt and gross wealth have negative effects on physicians' supply of community health services. Gross debt and gross wealth have no effect on fee-for-service income per listed person in the physician's practice, and positive effects on the total income from fee-for-service. The higher income from fee-for-service results from a longer patient list. Patient shortage has no significant effect on physicians' supply of community services, a positive effect on the fee-for-service income per listed person, and a negative effect on the total income from fee for service. The fact that the estimated effect of patient shortage is not statistically significant in Essay 4 might be a result of excluding wage rate in the community health service from the analysis. The reason for excluding this variable was that information on the community service reimbursement rates was of poorer quality in the second year of data collection. Another possible explanation is that the GPs expected future income and spending, as captured by the gross wealth and gross debt, is correlated with the indicator of patient shortage.

1.4 Policy implications

The significant effort of the Norwegian authorities to involve the inhabitants in the introduction of the list system in general practice has resulted in unique data for studying patients' revealed preferences for GPs. The results presented in this dissertation indicate that the implementation procedure was significant in determining the size and composition of GPs' lists. List composition, in terms of age and gender distribution, is affected by the preferences of the demand side, as the results support the hypothesis that consumers prefer GPs who resemble themselves with respect to observable characteristics. This conclusion has several policy im-

plications. First, it is beneficial if the population of general practitioners in the market resembles the population of patients. This result has further implications regarding the substitutability of GPs, and hence the preconditions for competition in the market for GPs. One or two available GPs in the market is perhaps not enough to promote competition: An elderly female might hesitate to switch to a young male GP starting a new practice. The regulator might thus want to ensure that the group of GPs available to accept new patients consist of male and female GPs of different age categories.

The limited opportunity to switch GP in many municipalities has received much attention in the public debate. Presently the Norwegian Ministry of Health and Care Services proposes a sizable increase in the number of GPs, and at the same time a reduction in the legal maximum number of patients that can be listed in a single practice. The latter initiative reflects a suspicion that the so-called "list barons", with more than 2000 persons listed, offer poor access and long waits. However, if quality of services indeed has a positive impact on the demand facing the GP as suggested by the results in Essay 2, one might argue that the reason "list barons" have long lists is that they offer services of high quality. If this is the case, then prohibiting lists longer than, say, 2000 people will result in, ceteris paribus, a reduction in the average quality of services in the market as patients are forced to leave the "list barons" in order to be listed with GPs offering services of lower quality.

An important issue in the public debate in Norway the last few years has been nursing home residents' limited accessibility to GPs' services. As of today the consensus appears to be that it is essential to achieve an increase in physician-hours within nursing homes. The results from Essay 3 are important in this regard, showing that GPs will respond to an increase in reimbursement rates and supply more hours. However the supply of hours appears to be inelastic with respect to wage. This implies that relying on the existing GPs to provide the necessary increase in physician hours will most likely be costly. Thus an idea might be to reconsider the strategy where physician hours in nursing homes are mainly provided by GPs in part time positions. Hiring fulltime physicians to provide physician hours in the nursing homes will also address the critique raised by the Norwegian Board of Health Supervision, namely that relying on part time positions in the nursing

homes results in a scattering of medical responsibility. According to current regulations, the municipalities can require GPs to work up to 7.5 hours per week with community health service tasks and some even argue that the general practitioners duties in the community health service should be increased to 15 hours per week. The results from Essay 4 support the idea of preference heterogeneity. If one accepts that there is heterogeneity in preferences, then working a certain number of hours in the community health service is likely to affect the physicians differently. The 7.5-hour minimum might be considered a type of "forced labor", implying that a gain in allocative efficiency of labor is likely to result if the minimum requirement were removed from the current regulations.

1.5 Limitations and suggestions for further research

The fact that we make no attempt to model the demand side for labor in community health service in Essays 3 and 4 might be considered to be a limitation of these studies. Further one might question the assumption that patient shortage is exogenous in the censored regression models in these essays. Including the demand side and modeling patient shortage as an endogenous event are ideas for further research.

Essays 1 and 2 are related, since they both consider the demand for GPs. One could note that they complement each other, each revealing limitations in the other: Essay 1 focuses mainly on what is often referred to as horizontal product differentiation (Coke versus Pepsi). The fact that people may choose different GPs for the same reason, for example to be listed with the closest GP or a GP having the same gender, is largely ignored in Essay 2. Similarly, the quality, in terms of vertical product differentiation (Lada versus Mercedes), is given less attention in Essay 1.

A topic addressed in both Essay 1 and Essay 2 is that individuals engaging in active provider choice may be different from individuals who choose a random provider or choose to accept the provider assigned by the authorities. As discussed in Essay 2, heterogeneity in individuals' propensity to participate in active provider choice may induce selection mechanisms resulting in systematic differences in patient morbidity and patient mortality between GPs with different levels of quality.

There are limits to the number of patients that can be listed with a GP. If one believe that high quality GPs experience higher demand than providers of lower quality, the high quality GPs are more likely to close the list in order to manage the workload. The result is that high quality providers are removed from the opportunity set of passive or "slow" decision makers. The result of this selection mechanism is not obvious. The results used to generate the estimation sample in Essay 1 suggest that individuals taking active part in choice processes have more education, higher wealth and income and are less likely to be unemployed. The results presented in Table 6 in Essay 2 indicate that individuals who took active part in provider choice have lower expected mortality. Viewing these results together, it is tempting to suggest that it was the healthiest individuals who took active part in the GP choice process, and further, that these individuals were therefore listed with the GPs offering the highest quality. The results presented in Essay 2, however, indicate that the effect of GP quality on aggregate health state of listed patients is not significant. An idea for future research is to explore further the association between individuals' health state and participation in provider choice. Understanding this mechanism is important, as evidence of systematic patient selection would indicate that extensions of consumer choice in health care have implications for equitable distribution of health care services.

1.6 References

- Aigner, D.J., Hsiao, C., Kapteyn, A., and Wansbeek, T. (1984): Latent Variable Models in Econometrics, in Griliches, Z. and M.D. Intriligator (eds.): *Handbook of econometrics, vol. II*, Amsterdam: North-Holland, 1321-1393.
- Arrow, K.E. (1963): Uncertainty and the welfare economics of medical care. *American Economic Review* **53**, 941-973.
- Culyer, A. J., and Newhouse, J. P. (2000): Introduction: the state and scope of health economics. In A. J. Culyer, and J. P. Newhouse (Eds.), Handbook of Health Economics. Amsterdam, North Holland.
- Dusheiko, M., Gravelle, H., Jacobs, R., and Smith, P. (2006): The effect of financial incentives on gatekeeping doctors: evidence from a natural experiment, *Journal of Health Economics*, **25**: 449-478.
- Fuchs, V.R., (1986): Physician-induced demand: a parable. Journal of Health Economics 5, 367.
- Fuchs, V.R., (2000): The future of health economics. Journal of Health Economics 19, 141-157.
- Gaynor, M., (2006): What Do We Know About Competition And Quality In Health Care Markets? NBER working paper No. w12301. National Bureau of Economic Research.

- Godager, G. T.Iversen and H. Lurås, (2007): Fastlegeordningen. Utvikling i bruk, tilgjengelighet og fornøydhet Working paper **2007:6** Health Economics Research Programme (HERO), University of Oslo.(In Norwegian)
- Gravelle H, M. Sutton and A. Ma, (2008): Doctor behaviour under a pay for performance contract: Further evidence from the quality and outcomes framework. CHE Research Paper 34.
- Iversen, T and H. Lurås, (2006): Capitation and incentives in primary care. In Jones, A. M. (ed.) The Elgar Companion to Health Economics, Cheltenham UK, Edward Elgar, p. 269-78
- Jones, A., (2000): Health econometrics. In: Culyer, A. J., Newhouse, J. P. (Eds.), Handbook of Health Economics. Elsevier, Amsterdam, North Holland
- Jöreskog, K.G., (1977): Structural Equation Models in the Social Sciences: Specification, Estimation and Testing, in P.R. Krishnaiah (ed.): *Applications of statistics*. Amsterdam: North Holland, pp. 265-287.
- Lurås, H., B. Paulsen, and J.E. Finnvold, (2003): De fleste fikk den de ville ha. (In Norwegian). Samfunnsspeilet 05/2003.
- Lurås, H., (2007): The association between patient shortage and patient satisfaction with general practitioners. Scandinavian Journal of Primary Health Care, 25, 133-139.
- McGuire, T.G., (2000): Physician agency. In A.J. Culyer and J.P. Newhouse (Eds.) *Handbook of Health Economics*, Volume 1, Elsevier Science, Amsterdam, 461-536.
- Moffitt R. A., (1999): New developments in econometric methods for labor market analysis. In Handbook of Labor Economics, Ashenfelter OC, Card D (eds). Elsevier: Amsterdam; 1367-1397.
- Newhouse Joseph P., (1996): Reimbursing Health Plans and Health Providers: Efficiency in Production Versus Selection *Journal of Economic Literature*, **34:3**, 1236-1263
- Popper, K. R., (1959): The logic of scientific discovery. London: Hutchinson.
- Scott, A., (2000): Economics of general practice. In A.J. Culyer and J.P. Newhouse (Eds.) *Hand-book of Health Economics*, Volume 1, Elsevier Science, Amsterdam, 1175-1200.
- Smith, M., (1998): Social science in question, SAGE Publications, London.
- Tobin J., 1958: Estimation of relationships for limited dependent variables. *Econometrica* **26**: 24-36.

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 P.O. Box 1089 Blindern, 0317 Oslo, Norway

> E-mail: geir.godager@medisin.uio.no Tele: (+47) 22 84 50 29 Fax: (+47) 22 84 50 91

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 \in 0.89 per consultation, whereas the estimated willingness to pay for having a GP with the same gender amounts to respectively \in 1.71 and \in 3.55 for female and male decision makers, respectively.

KEYWORDS: GP services. Discrete choice. Willingness-to-pay. Health care demand.

JEL CLASSIFICATION: C25, C83, D12, I11

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 calculate¹, 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

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

²The 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 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

³Confront 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			
	POPULATION	SELF-SELECTED	CORRECTED	
Variable	N=401999	N=15000	N=15000	
	Proportion	Proportion	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 *qenerate* 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...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

⁴Representative 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 \in 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

⁵We 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†	(€)/км	(€)/HOUR‡	
TAXI	(0) 4.10	1.30	6.80	TRAVELCOSTS = $2 \times [4.10 + 1.30 \text{ KM} + 6.80 \text{ HRS}]$
CAR	0	0.40	6.80	Travelcosts = $2 \times [0.40 \text{ km} + 6.80 \text{ hrs}]$

[†] 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 makers⁷. 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

⁶I 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. 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:

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

⁸This 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, ..., 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 X_{nj} and Z_j are vectors of explanatory variables and β and γ 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. X_{nj} are explanatory variables interacting characteristics of the GP j with characteristics of consumer n, and Z_j are explanatory variables describing characteristics or attributes of GP j. In contrast with X_{nj} , Z_j does not show any variation between decision makers, or in other words, there are no "within alternative" variation in Z_j . These vectors will include the following variables:

$$m{X}'_{nj} = \left[egin{array}{c} ext{GENDER}_{ff} \ ext{GENDER}_{mm} \ ext{AGEDIFFERENCE} \ ext{TRAVEL COSTS} \end{array}
ight], \, m{Z}'_j = \left[egin{array}{c} ext{SPECIALIST} \ ext{NORWEGIANGP} \ ext{MARRIEDGP} \ ext{AGEGP} \ ext{AGEGP} \ ext{AREA}_1 \ ext{} \vdots \ ext{AREA}_K \end{array}
ight]$$

We see that the \mathbf{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 $\mathbf{X}_{nj}, \mathbf{Z}_j$, 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:

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

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 regions⁹: 01, 02, 03, ..., 12. The postal code region are used to define the five areas referred to as WEST, NORTH, EAST₁, EAST₂, 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

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

¹⁰The 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 Trav	vel Model	Car Trav	el Model
Regressor	ESTIMATE	Std.Err.	ESTIMATE	Std.Err.
X_{nj} variables				
$\operatorname{GENDER}_{ff}$	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**
$oldsymbol{Z}_j$ 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*
EAST ₁	-0.47	0.07**	-0.33	0.07**
EAST ₂	-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**
EAST ₁	0.86	0.01**	0.73	0.01**
EAST ₂	0.78	0.02**	0.67	0.02**
SOUTH	0.88	0.02**	0.69	0.02**
Log likelihood	-68267.33		-6812	28.09
P-value LR test for IIA	Prob > chi	2 = 0.0000	Prob > chi	2 = 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 theory¹¹. 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 alter	RNATIVE: F	PRIVATE CAR		
Variable	WTP est. in €	95 % Conf. Int.		P est. in € 95 % Conf. Int.		WTP est. in €	95 % Co	ONF. 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

¹¹Dissimilarity 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:

$$\frac{\Delta \text{ TRAVEL COSTS}}{\Delta \text{ SPECIALIST}} = -\frac{\gamma}{\beta}.$$
 (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 \in 0.89 per consultation, whereas the estimated willingness to pay for having a GP with the same gender amounts to respectively \in 1.71 and \in 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 β and γ 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 the authorities' willing-

ness 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 \leq 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. McIntosh 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	POPULATION	RANDOM SAMPLE	CORRECTED SAMPLE
		Participants only	Participants only
Dummies	N=401999	N=15000	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

 $\label{eq:table A.2: Results from Logit Estimation} Table A.2: Results from Logit Estimation \\$ 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
	0.265	0.047 **
EDUCATION ₃		0.047
EDUCATION ₄	0.339	0.047
EDUCATION ₅	0.340	0.002
Education ₆	0.478	0.047
EDUCATION ₇	0.518	0.048
EDUCATION ₈	0.526	
Missingedu	-0.554	0.049 **
AGECAT ₂	0.229	0.010 **
AGECAT3	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 **
DECILE3	-0.376	0.018 **
DECILE ₄	-0.389	0.018 **
DECILE ₅	-0.427	0.019 **
DECILE ₅	-0.427	0.019 **
DECILE ₇	0.006	0.019 **
DECILE ₈	0.000	0.019 **
DECILE8	0.095	0.019 **
	0.090	0.019
Total Income	0.450	0.018 **
Decile ₁	-0.173	0.010
Decile ₂	-0.150	0.010
Decile ₃	0.005	0.018 **
Decile4	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 **
CITYAREA2	0.261	0.012
CITYAREA3	0.221	0.018 **
CITYAREA4	0.279	0.018 **
CITYAREA5	0.206	0.017 **
CITYAREA6	0.464	0.017 **
CITYAREA ₇	0.464	0.021 **
	$0.404 \\ 0.591$	0.021 **
CITYAREA8	0.391	0.024
CITYAREA9		0.013
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 $|\hat{\rho}_{np} - \hat{\rho}_p| < 0.01$

$\hat{ ho}_{np}$	# to replace	# matching candidates	$\hat{ ho}_{np}$	# 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.

DOES QUALITY INFLUENCE CHOICE OF GENERAL PRACTITIONER? AN ANALYSIS OF MATCHED DOCTOR-PATIENT PANEL DATA

ERIK BIØRN

Department of Economics, University of Oslo, P.O. Box 1095 Blindern, 0317 Oslo, Norway

E-mail: erik.biorn@econ.uio.no

GEIR GODAGER

Institute of Health Management and Health Economics, University of Oslo, P.O. Box 1089 Blindern, 0317 Oslo, Norway

E-mail: geir.godager@medisin.uio.no

ABSTRACT: The impact of quality on the demand facing health care providers has important implications for the industrial organization of health care markets. In this paper we study the consumers' choice of general practitioner (GP) assuming they are unable to observe the true quality of GP services. A panel data set for 484 Norwegian GPs, with summary information on their patient stocks, renders the opportunity to identify and measure the impact of GP quality on the demand, accounting for patient health heterogeneity in several ways. We apply modeling and estimation procedures involving latent structural variables, *inter alia*, a LISREL type of model, is used. The patient excess mortality rate at the GP level is one indicator of the quality. We estimate the effect of this quality variable on the demand for each GP's services. Our results, obtained from two different econometric model versions, indicate that GP quality has a clear positive effect on demand.

KEYWORDS: GP services. Health care quality. Health care demand. Latent variables. LISREL. Panel data. Norway

JEL CLASSIFICATION: C23, C33, D83, H51, H75, I11, I18

ACKNOWLEDGEMENTS: We are grateful for comments from Tor Iversen, Ismo Linnosmaa, Karin Monstad, Terje Skjerpen, Gilberto Turati, and participants at the 27th Nordic Health Economists Study Group Meeting, Copenhagen, August 2006, a Workshop on Health Economics, Oslo, August 2006, the 6th iHEA World Congress in Health Economics, Copenhagen, July 2007, and an Health Econometrics Workshop, Milano, December 2008. Godager acknowledges financial support from the Norwegian Ministry of Health and Care Services and the Research Council of Norway through the Health Economics Research Program at the University of Oslo (HERO). Some of the data applied are provided by the National Insurance Administration and Statistics Norway and have been prepared for research purposes by The Norwegian Social Science Data Services. The authors alone are responsible for all analyses and interpretations.

1 Introduction

Asymmetric information between physicians and their patients is a basic characteristic of the market for health care services. In the words of Arrow (1963):

"...medical knowledge is so complicated, the information possessed by the physician as to the consequences and possibilities of treatment is necessarily very much greater than that of the patient, or at least so it is believed by both parties".

Patients are therefore often considered to be poor judges of service quality. However, those who have repeated encounters with the same health care provider, will accumulate information on services and treatment outcomes, thus narrowing the information gap. The market for general practitioners' services is characterized by durable doctor-patient relations that may improve the patients' quality assessment. The aim of this paper is to investigate empirically whether the demand facing a general practitioner (GP) responds to the quality of the provided services.

General background

The impact of quality on the demand facing health care providers has important implications for the organization of health care markets. There is a growing literature on competition and quality in such markets, from which an important result is that the effect of stronger competition on quality depends crucially on the relative sizes of the price elasticity and the quality elasticity of demand. More competition may bring about reductions in quality if the quality elasticity is small compared to the price elasticity (Dranove and Satterthwaite, 2000, Gaynor, 2006). Further, the impact of quality on the demand facing health care providers has important implications for the optimal design of payment systems. A familiar result is that a retrospective payment scheme in the form of cost reimbursement is likely to pursue the goal of quality provision while giving weak incentives to provide cost reducing efforts. Conversely, prospective payment schemes tend to strengthen the incentives for cost reduction, while weakening the incentives for providing quality. A combination of payment mechanisms is thus likely to perform better than payment systems employing only one parameter. However, if quality affects demand, a first-best solution can, in theory, be obtained under a pure prospective payment scheme (Ma, 1994). This suggests that the effect of quality on demand – and information on its numerical size – is a key factor determining the optimal calibration of the parameters in the payment system: If the market punishes providers who are skimping on quality, the payment system can put more weight on the parameters that encourage cost reducing efforts.

Relation to literature

A conventional empirical approach when seeking to assess the effect of quality on demand for health care services is to estimate the effect of provider characteristics on individual consumers' choice of provider, applying different models for individuals' discrete choice. An influential paper in this tradition is Luft et al. (1990). They specifically study the effect of quality indicators such as death and complication rates, teaching status of hospital, and out of state admissions on patients' choice of hospital, using logit models, and find positive effects for several of the applied indicators. Using similar quality indicators and methods, Burns and Wholey (1992) extend the framework by including in their logit models characteristics of the admitting physician. They find that quality affects demand positively, and that characteristics of the admitting physician are important determinants of patients' hospital choice. More recently, Howard (2005), applying a mixed logit model on data on kidney transplantations, estimated the effect of the deviation from expected failure rate on probabilities of hospital choice. The results indicate that hospitals with a higher than expected failure rate have smaller probabilities for being chosen. A different empirical strategy is followed in Chirikos (1992), in estimating, by linear regression, the effect of individual hospitals' quality spending on their market shares. The results support the hypothesis that increased provider quality affects demand positively.

The present paper adds to the literature in several ways. First, no previous empirical studies seem to have considered the demand effects of quality in the market for general practitioners. Second, in the current literature the relationship between demand and various indicators of quality, such as mortality rates, failure rates or hospital type, and other independent variables, are estimated separately. The present paper contributes to the literature by simultaneously estimating the relationship between demand and quality and the way in which quality becomes manifest via indicators, applying linear structural equation modeling (LISREL) and estimation methods. Taking this approach, we acknowledge both the multidimensional aspect of the quality concept, and that it may be considered as more appropriate to interpret outcome measures such as mortality rates or failure rates as indicators of quality, rather than as direct measures of quality itself. Third, our econometric model provides a method to separate the effect of quality on outcome measures from the effect of patient health.

Setting of the study

In June 2001 a regular GP scheme was introduced in Norwegian general practice, making the GPs responsible for the provision of primary care services to the persons listed at their practice. Prior to the reform the health authorities gathered the information needed to assign one GP to each Norwegian inhabitant. All inhabitants were asked to rank their three most preferred GPs in a form, and all GPs were asked to report the maximum number of patients they would like to take care of. An algorithm was designed to utilize this information and obtain a one-to-one match between inhabitants and GPs.

Our data set has a panel format with the GP as the observation unit, but for some variables only one observation per GP exists. The data stem from The Norwegian General Practitioners Database, covering all Norwegian GPs, supplemented by measures of the

GP density in each municipality and of age-gender specific mortality rates. Among the variables recorded are the number of persons who ranked each GP as most preferred when returning the entry form, the number of mortalities among each GP's listed patients during a six-month period, and the proportion of the listed persons who switch to other GPs in later periods. For a stratified sample of GPs, relating to 14 municipalities, from this official GP database the data set has been extended to also include the median income and wealth of the listed persons and the proportion of them who have not finished high-school. In the analysis, we interpret the number of first-rankings and the proportion of listed persons who switch to other GPs, as indicators of the demand facing each GP. Our main hypothesis is that there exists a latent stochastic variable, denoted as GP quality, which, when heterogeneity related to the health status of the listed persons and other observed heterogeneity have been accounted for, is positively related to the demand facing each individual GP and negatively related to the recorded excess mortality of the GP's listed patients. We find empirical support to this hypothesis.

Two kinds of models are considered: a panel data model with latent heterogeneity related to GP quality and a multi-equation *LISREL type of model*, including both GP quality and the health of the stock of persons on the GP's list as latent variables, both of which are assumed to affect demand as well as other observed variables. For some variables, including the proportion of persons switching and the excess mortality, we have data in the panel data format. This is profitable for quantifying the latent heterogeneity and its consequences.

The rest of the paper proceeds as follows. The modeling of the demand in the market for GPs is discussed in the following two sections. In Section 2, we present a theoretical argument supporting the view that the expected demand facing each individual GP can be a function of quality, even if the true quality is unobserved to potential patients. This motivates testable predictions and hypotheses to be examined in the paper. In Section 3, we present the two econometric models. The data are described in Section 4, while estimation and test results are presented in Section 5. In Section 6 we discuss the results and conclude.

2 Some theory: Quality and demand

Since consumers are imperfectly informed about the quality of GPs, we distinguish between *true* and *perceived* quality. In order to understand how individuals' *perceptions* of quality of available providers can form the basis for the *demand* facing the providers we may turn to the literature on probabilistic choice models, cf. McFadden (1981).

Let μ_j denote the true quality of GP j (j = 1, ..., M), while q_{ij} denotes the quality of GP j as perceived by consumer i (i = 1, ..., N). The relation between true and perceived quality may be expressed as $q_{ij} = \mu_j + u_{ij}$ where u_{ij} denotes a noise term or measurement error. Conditional on μ_j , the probability that consumer n chooses GP i

can now be expressed as:

$$P_{ni} = P(\mu_i + u_{ni} > \mu_j + u_{nj}; \forall i \neq j) = P(u_{nj} < u_{ni} + \mu_i - \mu_j; \forall i \neq j)$$

Different assumptions for the noise terms u_{ij} lead to different types of choice models. Assuming that the u_{ij} terms are independent, and follow an extreme value type 1 distribution, will result in choice probabilities consistent with a logit type model. In that case we get the well known closed form expressions for P_{ni} , given by: $P_{ni} = \frac{e^{\mu_i}}{\sum_k e^{\mu_k}}$, cf. Train (2003, p. 78). In this simple model the demand facing provider i may be expressed as NP_{ni} . We have the following predictions from this simple theory:

[P1] GPs with high quality of services have a higher probability of being selected by a representative consumer than a GP whose services are of lower quality.

[P2] The selection probabilities P_{ni} are independent of the number of consumers, N. For a given population of M GPs, expected demand for the services of any of them, is a linear function of N.

This simple model may be generalized an extended in many ways. In particular, one may argue that consumers are able to affect the precision of their own quality assessment. If some consumers are more skillful or eager in gathering and processing information in the market than others, the result will be heteroskedasticity in the u_{ij} terms: Relaxing the homoskedasticity assumption of u_{ij} , denoting instead the variance of u_{ij} by σ_{ij}^2 , this could have been accounted for by allowing for $\sigma_{ij}^2 < \sigma_{hj}^2$ if consumer i has taken efforts to become better informed about GP j's quality than has consumer h. A prediction from such an extended model is that high-quality GPs tend to have a higher proportion of skilled or eager consumers on their lists than low-quality GPs. The possible existence of such a selection mechanism is important since consumers who are skilled or eager in collecting information, may have a health status and a death probability different from those not so skilled or eager. The crucial question then becomes: which groups of consumers, according to observable characteristics, devote most attention and efforts in searching for the best GP? On the one hand, less healthy consumers, with a high expected mortality rate, may be thought to be particularly concerned about their choice of GP and as a result be more willing to collect information than the average consumer. This may contribute to increasing the average mortality rate among the patients listed with highquality GPs. On the other hand, consumers who are more healthy and resourceful and have low expected mortality may be particularly able to collect and process such information. This may contribute to the outcome of the selection mechanism being reversed, *i.e.*, lowering the average mortality rates of the persons listed with high-quality GPs. Consequently, from a priori reasoning it is not obvious that the outcome of (observed or unobserved) patient heterogeneity will be neither that high-quality GPs attract patients with an average health status which differs from that of the low-quality GPs, nor if there is a difference, in which direction it will go. If a mechanism systematically selecting patients with different expected mortality rates for GPs of different professional quality is at work, and heterogeneity in health status among listed patients is given insufficient

attention in our modelling, we are likely to face severe difficulties when trying to estimate the impact on demand of GP service quality. The models to be described below have different degree of sophistication and are not equally well designed to meet this challenge. We address this issue in more detail in sections 3 and 5.

3 Econometric models

Motivation

In order to represent, and hopefully quantify, how the demand for GP services responds to GP quality and other relevant variables – as motivated by the theoretical argument put forth in the previous section – two kinds of models will be considered. The first, *Model A*, is a two-equation panel data model accounting for latent unit-specific heterogeneity. We associate the latter with, inter alia, perceived GP quality. The second, Model B, is a more complex, multi-equation model of the LISREL type. It includes not only GP quality among its latent variables, but also the initial health status of the persons entered on the GPs' lists. This extension serves to control for the fact that GP quality and observed GP heterogeneity may interact with observed and latent heterogeneity of the listed persons in multiple ways when determining demand as observed in the market.

Model A: Two-equation panel data random effects regression model

Assume that, in a certain district, at time t, there are M_t GPs, indexed by, $j=1,\ldots,M_t$, and N_t patients, indexed by $i=1,\ldots,N_t$. As before, we let μ_j denote the true quality of GP j, unobserved both to the consumers and the health administrators, and now treated as stochastic. Let further y_{1jt} and y_{2jt} denote two observable variables, which may be considered indicators of μ_j at time t. The interpretation adopted in Model A is that y_{1jt} is the demand facing GP j, and y_{2jt} is the excess death rate of persons on the list of this GP at time t. We specifically measure demand only by the number of consumers ranking the GP as the most strongly preferred prior to the implementation of the regular GP reform, and it is observed in period t=1 only. The variables assumed to explain (y_{1j1}, y_{2jt}) are quality and observable variables, of which some vary across both GPs and time periods, denoted as two-dimensional variables, and some are GP-specific.

We specify

$$(1) \qquad y_{1j1} = \boldsymbol{x}_{1j1}\boldsymbol{\beta}_{1} + \boldsymbol{z}_{1j}\boldsymbol{\gamma}_{1} + \alpha_{1j} + u_{1j1}, \quad j = 1, \dots, M_{1}, \\ y_{2jt} = \boldsymbol{x}_{2jt}\boldsymbol{\beta}_{2} + \boldsymbol{z}_{2j}\boldsymbol{\gamma}_{2} + \alpha_{2j} + u_{2jt}, \quad j = 1, \dots, M_{t}; \ t = 1, \dots, T, \\ (2) \qquad \left(\begin{bmatrix} u_{1j1} \\ u_{2jt} \end{bmatrix} \middle| \begin{bmatrix} \boldsymbol{x}_{1j1}, \boldsymbol{z}_{1j}, \alpha_{1j} \\ \boldsymbol{x}_{2jt}, \boldsymbol{z}_{2j}, \alpha_{2j} \end{bmatrix} \right) \sim \mathsf{IID}(\boldsymbol{0}, \boldsymbol{\Sigma}), \quad \boldsymbol{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \ \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{u1u1} & \sigma_{u1u2} \\ \sigma_{u1u2} & \sigma_{u2u2} \end{bmatrix},$$

where $(\boldsymbol{x}_{1j1}, \boldsymbol{x}_{2jt})$ and $(\boldsymbol{z}_{1j}, \boldsymbol{z}_{2j})$ are the row vectors of two-dimensional and GP-specific variables, respectively, $\boldsymbol{\beta}_1, \boldsymbol{\gamma}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}_2$ are column vectors of coefficients, and $(\alpha_{1j}, \alpha_{2j})$ are stochastic latent variables relating to the GP j's quality, the latter assumed to affect

patients' demand as well as their mortality. A crucial part of the model are the equations which connect these latent variables with the latent quality μ_j . We consider two ways of formalizing this relationship stochastically, denoted as Versions 1 and 2. In both versions, parallel with the extended scope of the model, the statistical status of μ_j will be changed from being a deterministic expectation, interpreted conditionally, to being a latent stochastic variable, the distribution of which is a specific part of the econometric panel data model.

LATENT HETEROGENEITY. VERSION 1: We first specify

(3)
$$\alpha_{1j} = \lambda_1 \mu_j + \varepsilon_{1j}, \\ \alpha_{2j} = \lambda_2 \mu_j + \varepsilon_{2j},$$

$$\left(\begin{bmatrix} \mu_j \\ \varepsilon_{1j} \\ \varepsilon_{2j} \end{bmatrix} \middle| \begin{bmatrix} \boldsymbol{x}_{1j1}, \boldsymbol{z}_{1j} \\ \boldsymbol{x}_{2jt}, \boldsymbol{z}_{2j} \end{bmatrix} \right) \sim \text{IID} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\mu}^2 & 0 & 0 \\ 0 & \sigma_{\varepsilon 1 \varepsilon 1} & \sigma_{\varepsilon 1 \varepsilon 2} \\ 0 & \sigma_{\varepsilon 2 \varepsilon 1} & \sigma_{\varepsilon 2 \varepsilon 2} \end{bmatrix} \right),$$

where we expect $\lambda_1 > 0$, $\lambda_2 < 0$, and $\sigma_{\varepsilon_1 \varepsilon_2} = \sigma_{\varepsilon_2 \varepsilon_1} < 0$. When μ_j is low, *i.e.*, when GP j is a low-quality doctor, then his/her patients will have a higher mortality rate than can be explained by $(\boldsymbol{x}_{2jt}, \boldsymbol{z}_{2j})$, and he/she will meet a lower demand than can be explained by $(\boldsymbol{x}_{1j1}, \boldsymbol{z}_{1j})$. Equations (1) and (3) define a four-equation system of structural equations explaining $(y_{1j1}, y_{2jt}, \alpha_{1j}, \alpha_{2j})$ by $(\boldsymbol{x}_{1j1}, \boldsymbol{x}_{2jt}, \boldsymbol{z}_{1j}, \boldsymbol{z}_{2j}, \mu_j)$ and noise terms. Inserting (3) into (1) yields the reduced form

(5)
$$y_{1j1} = \mathbf{x}_{1j1} \boldsymbol{\beta}_1 + \mathbf{z}_{1j} \boldsymbol{\gamma}_1 + \lambda_1 \mu_j + \varepsilon_{1j} + u_{1j1}, \quad j = 1, \dots, M_1, \\ y_{2jt} = \mathbf{x}_{2jt} \boldsymbol{\beta}_2 + \mathbf{z}_{2j} \boldsymbol{\gamma}_2 + \lambda_2 \mu_j + \varepsilon_{2j} + u_{2jt}, \quad j = 1, \dots, M_t; \ t = 1, \dots, T,$$

LATENT HETEROGENEITY. VERSION 2: The alternative version is

(6)
$$\alpha_{1j} = \lambda \alpha_{2j} + \varepsilon_j,$$

(7)
$$\left(\begin{bmatrix} \alpha_{2j} \\ \varepsilon_j \end{bmatrix} \middle| \begin{bmatrix} \boldsymbol{x}_{1j1}, \, \boldsymbol{z}_{1j} \\ \boldsymbol{x}_{2jt}, \, \boldsymbol{z}_{2j} \end{bmatrix} \right) \sim \mathsf{IID} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\alpha 2}^2 & 0 \\ 0 & \sigma_{\varepsilon}^2 \end{bmatrix} \right),$$

where we expect $\lambda < 0$. Equations (1) and (6) define a three-equation system of structural equations which explains $(y_{1j1}, y_{2jt}, \alpha_{1j})$ by $(\boldsymbol{x}_{1j1}, \boldsymbol{x}_{2jt}, \boldsymbol{z}_{1j}, \boldsymbol{z}_{2j}, \alpha_{2j})$ and noise terms. Inserting (6) into (1) we get, instead of (5), the reduced form

(8)
$$y_{1j1} = \mathbf{x}_{1j1} \boldsymbol{\beta}_1 + \mathbf{z}_{1j} \boldsymbol{\gamma}_1 + \lambda \alpha_{2j} + \varepsilon_j + u_{1j1}, \quad j = 1, \dots, M_1, \\ y_{2jt} = \mathbf{x}_{2jt} \boldsymbol{\beta}_2 + \mathbf{z}_{2j} \boldsymbol{\gamma}_2 + \alpha_{2j} + u_{2jt}, \qquad j = 1, \dots, M_t; \ t = 1, \dots, T.$$

The latter equations, with $\lambda = \lambda_1/\lambda_2$ and $\varepsilon_j = \varepsilon_{1j} - \lambda \varepsilon_{2j}$, could, of course, alternatively have been derived from (1) and (3). However, if (4) were assumed to hold, (8) would not be a reduced form, since α_{2j} is correlated with ε_{2j} and therefore with $\varepsilon_j + u_{1j1}$. The empirical implementation of Model A, to be presented in Section 5, relies on Version 2, in that estimation is done sequentially and a predicted value of α_{2j} obtained from the second equation in (8), the excess mortality equation, serves as a proxy for GP quality in the first equation, the demand equation. The estimators used in Section 5 will thus be consistent, but they would have been inconsistent if (4) were the valid stochastic process.

Model B: LISREL model with GP quality and patient health latent

Model A gives a rather restrictive, uni-directional description of how demand for GP services is related to GP quality. A LISREL model [see Goldberger (1972), Jöreskog (1977), Aigner *et al.* (1984, Sections 4 and 5), and Jöreskog *et al.* (2000)], may be a better solution to the problem of modeling sample separation. Model B belongs to this class.

Again, we exploit the panel design of our data set, with the GP as the observational unit, containing GP-specific time-series for some variables, including patient-switching and mortality rates, as well as GP-specific and patient specific time invariant variables. We let t be the time index and suppress the GP subscript. Boldface and slim letters denote matrices/vectors and scalars, respectively. Model B has three categories of variables: observable (manifest) structural variables, latent structural variables, and error/noise variables. The categorization is as follows:

Observable (manifest) structural variables:

```
y_1: Number of persons wanting to be entered on list initially, in period 1 (scalar)
```

 y_{2t} : Number of persons switching to another GP in period t (scalar)

 x_1 : Observed GP characteristics initially, in period 1 [(6×1)-vector]

 x_2 : Observed patient characteristics initially, in period 1 [(3×1)-vector]

 x_{3t} : Excess mortality of patient stock in period t (scalar)

 x_4 : Other time-invariant GP-characteristics unrelated to GP quality [(2×1)-vector]

$$\mathbf{y}_2 \equiv [y_{21}, \dots, y_{2T}]'$$

 $\mathbf{x}_3 \equiv [x_{31}, \dots, x_{3T}]'$

Latent structural variables:

 η : Demand directed towards GP (latent, time-invariant scalar)

 ξ_1 : GP quality (latent, time-invariant scalar)

 ξ_2 : Patient health (latent, time-invariant scalar)

Error/noise variables:

 ζ : Disturbance in demand function

 $\varepsilon_1, \varepsilon_{2t}$: Errors in the measurement equations for demand

 δ_1 : Errors in equations relating GP quality to GP characteristics [(6×1)-vector]

 δ_2 : Errors in equations relating patient health to patient characteristics. $[(3\times1)$ -vector]

 δ_{3t} : Errors in equations relating patient health and GP quality to excess mortality (scalar)

$$oldsymbol{\delta}_3 \equiv [\delta_{31}, \dots, \delta_{3T}]'$$
 $oldsymbol{arepsilon}_2 \equiv [arepsilon_{21}, \dots, arepsilon_{2T}]'$

A basic hypothesis of Model B is that GP quality, ξ_1 , and patient health status, ξ_2 are exogenous to the rest of the system. The quality variable ξ_1 corresponds to μ_j in Model A, Version 1. Time invariance and exogeneity are also assumed for the time invariant GP characteristics, x_4 , represented by the gender and the country of origin of the GP. These four variables are considered as determined from outside, inherent in the GP and in the patient, and hence not subjected to feedback from the rest of the system. This is an important assumption, which, for at least ξ_1 and ξ_2 , may be questioned. To some extent it will be modified later on (Section 5), in examining the robustness of the primary

conclusions concerning the link between GP quality and patient demand to changes in basic assumptions. These genuinely exogenous variables are, in Model B, indicated by observable 'counterparts', which, by assumption, become endogenous.

The model has four elements: (i) a demand function for GP services expressed in terms of latent variables, (ii) measurement equations indicating this latent demand, (iii) measurement equations indicating GP quality and health status of listed persons, and (iv) distributional assumptions for the latent exogenous variables and the error terms.

First, the *demand function*, relating latent demand (endogenous) to GP quality (exogenous) and latent health status and other characteristics of the listed persons (all exogenous), is:

(9)
$$\eta = \Gamma_{11}\xi_1 + \Gamma_{12}\xi_2 + \mathbf{\Gamma}_{14}\mathbf{x}_4 + \zeta = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} & \mathbf{\Gamma}_{14} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \mathbf{x}_4 \end{bmatrix} + \zeta.$$

We can interpret Γ_{11} , Γ_{12} , Γ_{14} as (vectors of) structural coefficients and ζ as a disturbance. Second, the measurement system for latent demand is

This subsystem expresses that $y_1, y_{21}, \ldots, y_{2T}$ are treated as T+1 observable indicators of the latent demand for GP services. Technically, in factor-analytic terminology, we can interpret Λ_{Y1} and Λ_{Y2} as factor loadings for, respectively, the number of persons wanting to be on the list initially (positive loading) and the number of persons switching to another GP in a later period (negative loading), on latent demand. In standard regression terminology, we can interpret Λ_{Y1} and Λ_{Y2} as the marginal effects of the latent variables on the corresponding observable variables. The error terms $(\varepsilon_1, \varepsilon_2)$ may contain measurement errors. Third, the measurement system for GP quality and patient health is specified as

(11)
$$\begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Lambda}_{X11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Lambda}_{X22} \\ \boldsymbol{\Lambda}_{X31} & \boldsymbol{\Lambda}_{X32} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \\ \boldsymbol{\delta}_3 \end{bmatrix}.$$

This subsystem expresses that the vector of observed GP characteristics, \boldsymbol{x}_1 , is related to latent GP quality, that the vector of observed patient characteristics is related to latent patient health, and that the T vector of excess mortalities, \boldsymbol{x}_3 , is related to both GP quality and patient health. Technically, in factor-analytic terminology, $\boldsymbol{\Lambda}_{X11}, \boldsymbol{\Lambda}_{X31}$ can be interpreted as, respectively, factor loadings for GP characteristics and excess patient mortality on latent GP quality. Likewise, $\boldsymbol{\Lambda}_{X22}, \boldsymbol{\Lambda}_{X32}$ can be interpreted as factor loadings for, respectively, patient characteristics and excess patient mortality on patient health. The error terms $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\delta}_3)$ may contain measurement errors.

Fourth, the process determining the latent exogenous variables ξ_1, ξ_2 is modeled in terms of their first-order and second-order moments as follows:

(12)
$$\mathsf{E} \left[\begin{array}{c} \xi_1 \\ \xi_2 \end{array} \right] = \left[\begin{array}{c} \mu_{\xi 1} \\ \mu_{\xi 2} \end{array} \right], \qquad \mathsf{V} \left[\begin{array}{c} \xi_1 \\ \xi_2 \end{array} \right] = \left[\begin{array}{c} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{array} \right],$$

while the distributions of the error and noise terms are assumed to satisfy

(13)
$$E[\zeta] = 0, \qquad V[\zeta] = \Psi,$$

$$(14) \qquad E\begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \qquad V\begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix} = \begin{bmatrix} \Theta_{\delta 11} & \Theta_{\delta 12} & \Theta_{\delta 13} \\ \Theta_{\delta 21} & \Theta_{\delta 22} & \Theta_{\delta 23} \\ \Theta_{\delta 31} & \Theta_{\delta 32} & \Theta_{\delta 33} \end{bmatrix},$$

$$(15) \qquad E\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{0} \end{bmatrix}, \qquad V\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} = \begin{bmatrix} \Theta_{\varepsilon 11} & \Theta_{\varepsilon 12} \\ \Theta_{\varepsilon 21} & \Theta_{\varepsilon 22} \end{bmatrix},$$

(16)
$$\begin{bmatrix} \xi_1 \\ \xi_2 \\ x_4 \end{bmatrix} \perp \zeta \perp \begin{bmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \end{bmatrix} \perp \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}.$$

The final assumption, (16), where \bot denotes orthogonal, expresses, inter alia, the assumed exogeneity for GP quality and patient health. Its essence is that these variables, being modeled by (12), remain unaffected by the perturbations in the demand equation disturbances, and the errors in the measurement systems for demand (endogenous) and latent GP quality and latent patient health (exogenous). Since arguments may be raised that this model disregards a possible effect of GP quality on the listed patients' initial health status, we will in addition consider a modified version, Model C, in which this potential link is modeled and hence may be tested for.

4 Data

Data sources and data design

Prior to the introduction of the regular GP scheme in June 2001, the health authorities gathered the information needed to assign GPs to the entire Norwegian population. All inhabitants were asked to rank their three most preferred GPs in an entry form. The GPs were asked to report the maximum number of patients they would like to take care of. The health authorities utilized this information as an input in an algorithm allocating inhabitants to GPs. Most people got listed with the GP whom they had consulted prior to the reform (Lurås, et al., 2003).

Our data stem from The Norwegian General Practitioners Database supplemented by a measure of the GP density, as calculated from the number of contracted GPs in each municipality in June 2001, as well as aggregate age/gender specific mortality rates. The latter are calculated by means of aggregate mortality rates constructed by Statistics Norway. The Norwegian General Practitioners Database contains information on all Norwegian GPs, and the variables describing the individual GPs practice is provided by the National Insurance Administration (NIA) every six month. The database is administered by the Norwegian Social Science Data Services, who merge the information reported by NIA with socio-demographic variables as income, wealth and marital status, registered by statistics Norway. For GPs practicing in 14 municipalities, sampled by stratification, the database also includes characteristics for the patients who were listed in the GP's practice in June 2001, such as the median income and median wealth, and the proportion who have not finished high-school. For each GP we know the number of persons who ranked the GP at the top when returning the entry form, in this paper to be given the interpretation as an indicator of the demand facing the GP. After the reform was implemented, the GP database is updated at regular intervals to give the number of persons who are actually listed in the practice. After excluding observations with key variables missing, our unbalanced panel data set consists of a sample of 484 GPs observed up to 7 six-month periods. The pattern of observation is described in Table 1, from which we see that 441, or 91 %, of the GPs are observed in all 7 periods.

Table 2 lists and defines the variables applied in this paper, Table 3 gives overall descriptive statistics for the variables, and Table 4 gives descriptive statistics of the GP-specific means of the time varying variables. Descriptive statistics for variables at the level of the municipality are given in Table 5. We distinguish between variables observed at the GP level and variables which are observed at the municipality level and hence are common to all GPs practising in the same municipality.

The symbols used for the observable variables in the exposition of Models A and B above, (x, y, z), have their empirical counterparts among the variables in Table 2. This correspondence is given below (the GP subscript, for simplicity, suppressed):

Model A:

$$y_1' = [DEMAND], \ y_{2t}' = [ACTMORT_t], \ \boldsymbol{x}_1' \ \text{is empty}, \ \boldsymbol{x}_{2t}' = [EXPMORT_t]$$

$$\boldsymbol{z}_{1}' = \begin{bmatrix} GPDENS \\ MARRIEDGP \\ SPECGEN \\ SPECCOM \\ SPECOTH \\ ALPHA \\ IMMIGRGP \\ FEMALEGP \\ AGEGPSQ \end{bmatrix}, \ \boldsymbol{z}_{2}' = \begin{bmatrix} CENTRAL \\ LESSCENT \\ LEASTCENT \\ LOEDUC \\ PINCOME \\ PWEALTH \\ SPECGEN \\ SPECCOM \\ SPECOTH \\ FEMALEGP \\ AGEGP \\ AGEGPSQ \end{bmatrix}$$

¹The GPs from the municipality Tromsø, 44 in total, were excluded from the sample. Here, the regular GP scheme was implemented already in 1993 and very few inhabitants returned the entry form.

Model B:

Variables at the GP level, including patient stock characteristics

The variables collected at the GP level and related to the mortality of the persons on the GP's list are DEAD, EXPDEAD, ACTMORT, EXPMORT, and EXCMORT. The number of individuals leaving the list and the number of mortalities on each individual GP's list during a six-month period is registered in the GP database, except for the year 2002, where this information is registered for the whole calendar year only. We have allocated the mortalities and the switches in this year on the two half-years, according to the list sizes in the two half-years. DEAD denotes the number of mortalities during a period, and ACTMORT measures mortality per thousand listed patients.

GPs with a relatively high proportion of elderly people on their lists are presumably recorded with a relatively high mortality rate in any period. In order to compensate for this heterogeneity when measuring excess mortality, we proceeded as follows: Among the information registered in the GP database is the number of listed males and females belonging to each of the age categories 0–7, 8–19, 20–29, 30–39, 40–49, 50–59, 60–69, 70–79 years, and 80 years and above. By applying the gender and the age specific death probabilities (Statistics Norway, 2005a) and the age distribution in Norway (Statistics Norway, 2005b), we can for each GP calculate the expected number of mortalities, EXPDEAD, and the expected per thousand mortality rate, EXPMORT, *i.e.*, EXPDEAD per thousand listed persons. This enables us to calculate EXCMORT: the difference between the actual and the expected mortality rates at the GP level, henceforth to be referred to as the GP-specific excess mortality rate

From Table 3 we see that the overall mean of the actual number of mortalities during a six-month period (DEAD) is 5.63, and from Table 4 that its *GP-specific mean* ranges from 0 to 22 mortalities. The overall mean of the expected number of mortalities (EXPDEAD) is 5.12, with range from 0.05 to 23.9. By combining the aggregate death probabilities and the age-gender distribution of listed patients we have obtained a two-dimensional variable, with a mean value not far from the actual mean number of deaths according to the mortality statistics included in the General Practitioners Database.

The overall mean number of mortalities per thousand listed persons is 4.76. Its GP-specific means range from zero to 28. EXPMORT has an overall mean of 4.24 deaths per thousand, and the GP-specific means ranges from 0.45 to 12.45. An important variable in the analyses is the excess mortality rate, EXCMORT. As explained, a positive (negative) value means that the mortality rate at the GP level is higher (lower) than expected from the age and gender distribution of the persons on each GP's list. We note that the overall mean of EXCMORT is positive. The reason for this could be, on the one hand that the mortality tables are constructed from cross-sectional variation in mortalities during a period of only one year, on the other hand that life expectancy is known to be lower than the national average in the municipality Oslo (Statistics Norway, 2006a & 2006b), which is the location of 426 of the 484 GPs in the data set. A third explanation may be that the number of deaths in the first period is somewhat overestimated.²

We denote the number of consumers ranking a specific GP as most preferred as DEMAND directed towards this GP. It is time-invariant, as the matching of GPs and patients has been undertaken only once, in 2001, when implementing the regular GP reform. The average GP was preferred by 826 inhabitants, but there is a lot of variation. The most popular GP in our data set was preferred by 3152 inhabitants, whereas some GPs were not preferred by any. Being requested by a large number of inhabitants in a municipality with a high GP density is not equivalent to be strongly requested in a municipality where the GP density is low. We have considered two ways of controlling for differences in GP density across municipalities. The primary one is to include a measure of GP density as an additional explanatory variable representing observed heterogeneity. The secondary one is to weight DEMAND by a measure of GP density. The specific measure of the GP density applied here is the number of GPs per thousand inhabitants in the municipality (GPDENSITY); see the next sub-section for an elaboration. The specific measure of weighted demand we used is the variables DEMAND1, obtained as the product of DEMAND and GPDENSITY. Taking DEMAND1 as the relevant demand variable in our analysis implies that a given number of first rankings is interpreted as a higher demand in a municipality with a high GP density than in a municipality where GP density is low. Prediction [P2] provides the rationale for transforming the demand variable in this way.

We have information on the number of persons leaving a GP's list in order to enroll on a competitor's list.³ We refer to the proportion of listed persons leaving the list and

²The period-specific means of ACTMORT seem to be higher in the first period than in the later periods. This is not unexpected as the first period is one month longer, a fact we have adjusted for simply by multiplying the number of mortalities in the first period with $\frac{6}{7}$. We suspect that mortalities in the period April to June 2001 may also have been registered, although with a lag. In estimating the models, we have therefore alternatively applied an adjustment factor of $\frac{6}{9}$, which would have been correct if mortalities from April to June 2001 were indeed included among those registered for the second half-year 2001. The main results are not affected by this modification of the adjustment factor.

³Patients leaving the list because they migrate or move to another municipality are excluded from these numbers. We thus interpret LEAKRATE as the proportion of the listed persons who switch because they actually prefer another GP.

switching to other GPs in the municipality as LEAKRATE, and interpret this variable as a time-varying indicator of the demand. Its overall mean is 3%, its GP-specific means vary from close to zero to 27%, and the between GP variation accounts for as much as 90% of the total variation.

Our data set also reports the GP's age (AGEGP), gender (FEMALEGP), and marital status (MARRIEDGP) as well as the GP's birth country. We see that the average GP is 47 years old, that 36 % of the GPs are females and that 69 % are married. We have constructed a binary variable, IMMIGRGP, equal to 1 if the GP is born in a non-Scandinavian country. About 5 % of the GPs in our sample have this property. The variable denoted salary contract when practicing as a GP, and we see that 5 % of the GPs have this kind of contract.

The number of patients *actually* listed in the practice at the beginning and end of each period is registered in the GP database. To take account of within-period changes of this variable, we construct the average of the numbers recorded at the beginning and at the end of each period, giving the variable LISTSIZE. Its overall mean is 1211 persons, while its GP-specific means range from 153 to 2620.

Our data set also reports whether the GP is a specialist in general medicine (dummy variable SPECGEN), in community medicine (dummy variable SPECCOM) or in another medical field (dummy variable SPECOTH) – all of which are time-varying, but the within-GP variation is small. Overall, 56% of the GPs are specialists in general medicine, 7% are specialists in community medicine. and 3% are specialists in an other field.

Our GP-level data also contain the following information on the patients who were listed in the practice in June 2001: the median net income and median net wealth among the listed patients older than 30 years, the proportion of listed patients who are older than 30 and have not finished high school and the proportion of the listed patients who submitted the entry form signalling GP preferences. By construction, these variables are uni-dimensional, as this information is not updated after the implementation of the regular General Practitioner Scheme. The income and wealth variables, measured in 1.000 NOK, pincome and pwealth, have overall means 196.5 and 63.4, respectively. Not unexpectedly, they vary considerably: the GP whose listed patients are on average richest, have a median income twice that of the GP whose listed patients have the lowest median income. The corresponding median wealth, pwealth extends from -94.2 to 467.3. In the LISREL analysis, after some trial runs, we decided to exclude pwealth from the variable list, in order to ensure convergence. We suspect that the reason for this is that income and wealth are highly correlated.

Finally, PFORMSUB denotes the proportion of the listed patients who submitted the entry form prior to the implementation of the regular General Practitioner Scheme. This variable varies from nearly zero to one, indicating that some GPs were not assigned any patients who submitted the entry form, while other GPs were assigned only patients who submitted the form. The mean of this variable is 0.73, indicating that the average GP

have a list where 73 % of the patients submitted the entry form. We denote by PEDUC the proportion of the listed patients who are older than 30 and have not finished high school. We see that the average GP have 45 % of the listed patients characterized by not having finished high school. This share also varies considerably, from 2 % to 81 %.

Variables at the level of the municipality

Some important variables reported are specific to the municipality in which the practice of the GP is localized. Statistics Norway has constructed an indicator of *centrality*, placing each Norwegian municipality in one of 4 centrality categories. This indicator captures, *inter alia*, the population density and the distance to the nearest city of a certain size. We refer to these categories, in an order of increasing centrality, as least central (LEASTCENT), less central (LESSCENT), central (CENTRAL) and most central (MOSTCENT). In our sample six municipalities are categorized as least central municipalities, one as being less central, three municipalities as being central, while three municipalities are categorized as most central. A description of the GP density measure, GPDENSITY, and the number of GPs within each municipality are given in Table 5. We see that the range of GPDENSITY is from 0.55 to 1.57 GPs per thousand inhabitants.

5 Estimation and test results

Our data set include several variables expected to be related to the individual quality of the GP, primarily the demand variable DEMAND and excess mortality variable EXCMORT. It is important to note that information on the number of patient mortalities at the level of the GP is not publicly available in Norway. It is thus highly unlikely that individuals' choice of GP is directly related to these numbers. We derived from our theoretical model in Section 2, the prediction, [P1], that a positive relationship exists between the quality of the individual GP and the demand facing the GP, even when individual consumers may have incorrect perceptions of GP quality. How perceptions are formed is unknown and we make no attempt to open this "black box", as an enquiry into the formation of human perceptions is beyond the scope of this paper. We let consumers be heterogeneous with regard to their preferences, the information they possess, and the way they process information. Consumers may choose the same GP for different reasons, or different GPs for the same reason. We expect however that the GPs' appearance, experiences from earlier consultations with available GPs, advice from relatives and friends and even rumor to enter the "black box" as inputs in the formation of individuals' quality perceptions.

As explained in Sections 3 and 4, the statistical modeling of the 'causality chain' giving rise to this relationship is rather different in the two econometric models we consider, Model A and Model B. In addition, mainly as a robustness check of our main conclusion, results obtained from a third model, Model C – essentially a modification of Model B in one important respect – will be briefly considered at the end.

The estimation procedure for **Model A**, Version 2, represented by Equation (8), is, as explained in Section 3, a stepwise procedure. Using in both steps modules in the STATA 9 software, we estimate in the first step the effect of the variables representing observed heterogeneity and other assumed exogenous variables on the mortality rates and extract the predicted value of the random effect for each GP in the sample, ALPHAHAT.⁴ In estimating this mortality equation, *i.e.*, the second equation of (8), we allow for the possibility that the residuals are not independent within municipalities, and report robust standard errors. In the second step, this prediction, treated as an exogenous variable, is inserted in the demand equation, *i.e.*, the first equation of (8).⁵

For the multi-equation model, **Model B**, we apply the Maximum Likelihood (ML) procedure in the LISREL 8.80 software.⁶ The actual number of linear equations to be simultaneously estimated is 25. In this model both the quality of the individual GPs and the unobserved aggregate health status of the listed patients occur as latent exogenous variables, as explained in Section 3.

Results for Model A

The mortality equation

The dependent variable in this equation is ACTMORT. Observable heterogeneity is controlled for in various ways. First, to control for differences in the age and gender distribution of the GPs' listed patients, we include the expected mortality rates EXPMORT as an explanatory variable. Second, to account for heterogeneity between municipalities of different centrality, we include the centrality dummies LEASTCENT, LESSCENT and CENTRAL. Third, to account for observable GP heterogeneity we include AGEGP, three GP speciality dummies as well as FEMALEGP. Fourth, we include variables describing the listed patients, with intention to control for the possibility that the average health status of patients varies between GPs. Since there is evidence in the medical literature that life expectancy and health status is related to education, income and wealth (Lantz et al., 1998, Papas et al., 1993), we include PEDUC, PINCOME and PWEALTH as proxies for the average health status of the persons listed with each GP. Fifth, as discussed in Section 2, still another kind of heterogeneity may also occur: individuals who chose not to submit the entry form stating their preferences for certain GPs, may have an average health

⁴See Hsiao (2003, Section 6.2.2.c) and Lee and Griffiths (1979) on the prediction of random effects from panel data.

⁵An even simpler alternative also considered is a single-equation model where the excess mortality rate is inserted directly as a quality indicator in the demand equation – in a sense merging the two equations in (8) into one equation. The underlying assumption is that the GP specific level of patient excess mortality is exogenous. This approach, however, is defective to the extent that the GPs have an inhomogeneous patient stock with respect to the average health status, which will induce a bias in the coefficient estimate of the quality variable. The results from this 'single-equation version' of Model A is presented in Appendix A, Table A.1

⁶The Covariances and Means to be analyzed are estimated by the EM procedure, as there are some missing observations due to the unbalanced panel data.

status different from those who returned the entry form. We take account of this by including PFORMSUB as an explanatory variable. Since the range of PEDUC and PFORMSUB is restricted to the (0,1) interval, we transform them by the log-odds using the formula $\ln(\frac{x}{1-x})$ in order to extend their range to $(-\infty, +\infty)$ which gives a better balance with the unbounded range of the other explanatory variables.

The results from the mortality equation are presented in Table 6. All of the estimated coefficients except LEASTCENT, SPECGEN, AGEGP and AGEGPSQ are statistically significant, and the overall R^2 is rather high: 0.5071. The coefficient of EXPMORT is positive, as expected, since high expected mortality should have a positive effect on actual mortality. Further, GPs having their practice in a central or less central municipality have a significantly lower patient mortality rate than GPs in most central municipalities. The negative estimated coefficient of LOSUBMIT indicates that GPs who were assigned a high proportion of the persons who expressed their GP preferences in advance, have lower patient mortality than GPs who obtained a low proportion of patients who actively selected their GP. The coefficient on the education variable is negative, which is not in accordance with intuition saying that GPs with a high proportion of low-educated patients have a higher mortality rate and may be attributed to education being correlated with income and wealth. The estimated coefficients of income and wealth have the expected negative signs. Since these variables are measured in 1000 NOK, an increase in PINCOME and PWEALTH of NOK 100.000 would be accompanied by a reduction in the mortality rates of 0.97 and 0.94 deaths per thousand, respectively. We see that, ceteris paribus, the patient mortality rate of GPs who are specialists in community medicine is significantly higher and that of GPs who are specialists in a field other than general medicine and community medicine is significantly lower than the patient mortality rate of other GPs. Finally, female GPs have, ceteris paribus, a lower mortality rate of their patient stock than male GPs.

Statistics describing the predicted values of the GP specific heterogeneity variables, i.e., of the α_{2j} s obtained from the second equation of (8), denoted as ALPHAHAT, are given at the head of Table 7. According to our interpretation of Model A, α_{2j} represents a latent variable that is linearly related to quality, confer Equations (3) and (6). A histogram is given in Figure 1. Its form is not very far from a bell-shape, although with an outlier at the right end, equal to 13.11 deaths per thousand.

We next proceed to consider the results for the demand equations, in which the AL-PHAHAT predictions are among its explanatory variables.

The demand equation

The estimation results for the demand equation, in the two versions explained above, are presented in Table 7. In panel A, to which we will give most attention, the dependent

⁷This intuition is supported by supplementary regressions in which income and wealth were excluded as regressors, giving a significantly positive coefficient of LOEDUC.

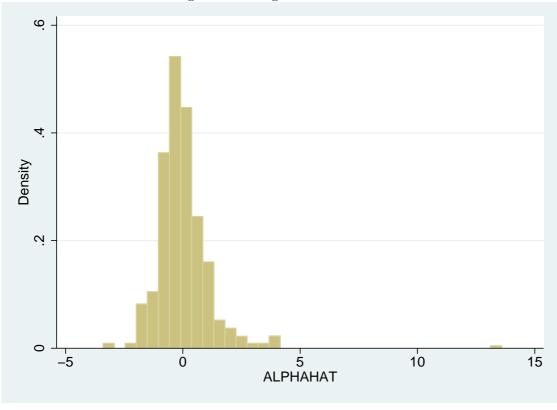


Figure 1: Histogram of ALPHAHAT

variable is DEMAND. Supplementary results, to investigate the sensitivity of the findings, when applying the GP density weighted measure of demand, DEMAND1, are given in panel B. In panel A, the GP density is included as a explanatory variable instead.

Explanatory variables included in both versions of the equation are the specialization dummies, SPECGEN and SPECOTH, which may be interpreted by the market as observable quality indicators. We also include the specialization dummy SPECCOM. We expect this variable to have a negative effect on demand. One reason for this may be that GPs who are specialists in community medicine are known to participate more frequently in the community health service than GPs who do not have this specialization (Godager and Lurås, 2007), and as a result, they may supply fewer business hours per week and hence appear less attractive for patients. To take account of observable GP heterogeneity we in addition include dummy variables MARRIEDGP, IMMIGRGP and FEMALEGP, as well as AGEGP as explanatory variables. Because the demand variable is time invariant, cf. the first equation of (8), weighted between-GP estimation is used, the weighting being motivated from the differing number of observations behind the GP-specific means.

The GP density has a statistically significant effect on demand, and its coefficient has the expected negative sign (Table 7, Panel A). Second, the estimated coefficient of ALPHAHAT is negative and statistically significant – supporting the intuition that increased

quality induces increased demand facing the GP. Furthermore, the estimated effect of MARRIEDGP and SPECGEN indicate that being married and being a specialist in general medicine contribute, ceteris paribus, to a higher market demand. While FEMALEGP is not statistically significant, AGEGP comes out with a positive and statistically significant effect. The latter results may be explained by the fact that a GP's age is correlated with the number of years in GP practice, and that a GP who has been practicing for a long time may be included in the choice set of a larger proportion of the consumers in the market. This mechanism, however, is not explicitly accounted for in our theoretical model. It would have been possible to capture it by introducing heterogeneous groups of consumers with different choice sets within the same market, and this can be done simply by furnishing the parameter N with a group subscript.

As to the signs of the effects as well as their significance, the results in Panel B are very similar to those in Panel A. Being married and being a specialist in general medicine are both estimated to have a positive effect according to this model as well, and again, AGEGP comes out with a significantly positive coefficient.

Results for Model B

A stylized path diagram corresponding to Model B, with some details (inter alia, the noise elements) omitted to preserve clarity, is given in Figure 2. The ovals and boxes indicate latent and observable variables, respectively. This diagram may be a useful reference for interpreting the results. The fact that no arrows point towards Quality and Health symbolizes their exogeneity. The variables having arrows directed towards them – including all observed variables except x_4 (not indicated in the figure, cf. Equations (9) and (16)) – are formally endogenous and correlated with at least one noise element. The measurement model specifies how Quality, Health and Demand (all latent) are indicated by observed variables; cf. Equations (11) and (10). As explained in Section 3, the model consists of measurement equations and structural equations, the latter representing the relationship between the three latent variables. In the sequel, we will let latent variables be indicated by names having capitalized first letter and let observable variables have non-capitalized first letter – following the conventional LISREL notation.

By modeling the demand facing the individual GP as a latent variable we are able to utilize information on the rate at which patients leave the GP's list in order to join a competitor's list. This approach thus takes into account how the demand facing the GPs has developed after the introduction of the General Practitioner Scheme. In our case the measurement model consists of two parts. The measurement equations for the exogenous latent variables, in the LISREL notation in Section 3 referred to as ξ -variables, are henceforth referred to as the X-measurement model. The measurement equations for the dependent latent variable, in LISREL notation referred to as an η -variable, is henceforth referred to as the Y-measurement model. When interpreting the approach and the results below, it should be recalled that the panel structure of the data – including the repeated

observations of patients leaving the GP's list as well as of the excess mortality rates at the level of the GP – is essential for obtaining the inference we want to make, as the three latent variables in focus on Model B, *Quality*, *Health* and *Demand*, are all time invariant.⁸ In total, Model B comprises 25 equations that are simultaneously estimated by the Maximum Likelihood (ML) method. We start by presenting the results from estimation of the structural equation before presenting the results from the estimation of the equations in the two measurement models.

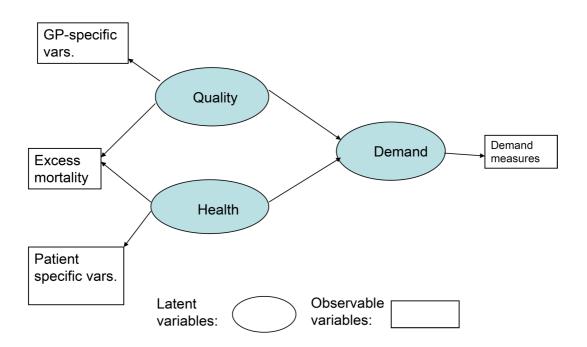


Figure 2: Model B: Stylized LISREL path diagram

The structural model (demand equation)

The results from estimation of the structural equation, corresponding to Equation (9), are given in Table 8.

Our first major finding is that the estimated effect of *Quality* on *Demand* is statistically significant and has the expected sign. Since the measurement scale of *Demand* is the number of first rankings measured in hundreds, the interpretation of the coefficient estimate of 1.36 obtained is that a marginal increase in *Quality*, equivalent to a marginal reduction in the excess mortality rates, increases *Demand* by 136 persons. Second, we

 $^{^8}$ As explained in Section 4, the observations from period 2 and 3 are linearly dependent by construction. Therefore observations from period 2 are excluded from the data set when estimating Model B.

find that the estimated coefficient of *Health* also is statistically significant with a positive sign, which implies that the more healthy a GP's patients are, for a given level of quality, the higher demand will be experienced. Also, the estimated coefficient of *femalegp* is negative and statistically significant.

As remarked in Section 2, heterogeneity of patients' ability to collect and process available quality information may induce selection mechanisms resulting in systematic differences in morbidity and mortality between GPs with different levels of quality. We have no prior knowledge of the direction in which this selection mechanism may go. However, failing to control for differing aggregate health status of listed patients is likely to result in a simultaneity bias and/or an excluded variable bias when estimating the effect of quality on demand. Our approach separates the effect of quality on excess mortalities from the effect of health at the level of the GP through the exclusion restrictions imposed on the measurement equations for Quality and Health: six variables describing the GP are included in the measurement equations for Quality, but excluded from the measurement equations for Health, while three variables are included in the measurement equations for Health but excluded from the measurement equations for Quality.

The estimated covariance matrix of the latent variables is given in Table 8, panel B. Our results indicate that there is a negative association between Quality and Health. What we observe is thus consistent with a situation where a selection mechanism exists such that GPs with low quality of services are endowed with a patient stock with a better health, as compared to GPs with higher quality of services. One may argue that Model B does not reveal the effect of the GPs' quality on the initial health state of listed patients, as it is set up to measure the effect of quality when controlling for initial health status of patients showing between GP variation. To address the latter issue, and for the purpose of conducting a robustness check of our main findings, we have additionally estimated an alternative LISREL model where all latent variables enter as η variables, i.e., as formally endogenous – Quality and Health being exogenous latent variables in Model B. Such a model setup allows the estimation of the marginal effect of GP's quality on the initial state of health. The results from this model, denoted as Model C, are reported in Appendix D. The results confirm the results from Model B, that latent quality affects latent demand positively. The numerical size of the effect is somewhat smaller, however. The most important single result from estimation of Model C is that quality is found not to have significant effect on the aggregate health status of the GP's listed patients.

6 Discussion and conclusion

Our analysis supports the hypothesis that the demand responds to the individual quality of the GP. Even though our two different approaches to quantifying this effect rely on different assumptions and different methods, it may be argued that our two sets of results pointing in the same direction, contribute to strengthening this conclusion. Our results indicate that a marginal quality increase equivalent to a reduction of the mortality rate

of one per thousand – being the implicit measurement scale of our quality indicator – increases the demand facing the GP by 57 persons according to Model A and by 136 persons according to Model B. In interpreting this finding one should recall that the empirical results for Model A (presented in Section 5), rely on Version 2 from Section 3, which imposes an asymmetry in the way quality affects latent GP-specific heterogeneity in the two equations. The estimators used under Model A, although enjoying consistency under Version 2, are inconsistent if the less restrictive Version 1 is valid.

The latter has been our primary motivation for also modeling latent heterogeneity within a LISREL framework. Our LISREL model, Model B, may be considered less restrictive than Model A and also enables us to address more appropriately the issue that the aggregate health status of listed patients, also considered an unobservable variable, is likely to be related to the mortality rates at the level of the GP.

We believe that this econometric modeling tool has a wider application in assessing the impact of quality of health care providers on demand for health services than the one presented here. Our LISREL model separates the effect of quality on outcome measures from the effect of patient health at the provider level through the exclusion restrictions imposed on the measurement equations for the latent variables Quality and Health. An idea for future research is to apply this modelling and estimation strategy to matched hospital-patient data, in order to measure the impact of hospital quality on demand. An important question from the point of view of policy implications is: how strong is the effect of quality on demand? In the context of regression models answers to such questions are often provided by means of appropriate elasticities. Due to the fact that the mean value of latent variables are undefined, the elasticities of interest, such as Quality elasticity of demand are also undefined. However, to get a 'metric' for assessing the magnitude of the effect we may compare the effect of quality on demand with the effect of the observable GP specific variables like the GP gender dummy and the dummy for whether or not the GP is born in a non-scandinavian country. In case of model B, this can be achieved by utilizing the standardized solution of our LISREL estimation in which all effects are scaled in terms of their standard deviation. The standardized solution thus obtained indicates that the effect of quality on demand is more than four times the effect of the GP gender dummy and ten times the effect of non-Scandinavian GP dummy. It is also illuminating to interpret the latter effect in relation to the standard deviation of the Quality variable σ_Q (confer Table 8), panel B. We then find that an increase in quality equivalent to a reduction of the patient mortality rate by one per thousand corresponds to a change in Quality equal to $0.6397\sigma_Q$. An increase of this order of magnitude results in an increase in Demand corresponding to 16% relative to the global mean of DEMAND, given in table Table 3.

References

- Arrow, K.E. 1963. Uncertainty and the welfare economics of medical care. *American Economic Review* **53**, 941-973.
- Aigner, D.J., Hsiao, C., Kapteyn, A., and Wansbeek, T. (1984): Latent Variable Models in Econometrics, in Griliches, Z. and M.D. Intriligator (eds.): *Handbook of econometrics, vol. II*, Amsterdam: North-Holland, 1321-1393.
- Burns, L.R. and Wholey, D.R. (1992): The Impact of Physician Characteristics in Conditional Choice Models for Hospital Care. *Journal of Health Economics* 11, 43-62.
- Chirikos, Thomas N.(1992): Quality Competition in Local Hospital Markets: Some Econometric Evidence from the Period 1982-1988. Social Science & Medicine 34, 1011-1021.
- Dranove, D. and Satterthwaite, M.A., (2000): The Industrial Organization Of Health Care Markets, in A.J. Culyer and J.P. Newhouse (eds.): *Handbook of Health Economics, vol.1*. Amsterdam: Elsevier Science, pp. 1093-1139.
- Gaynor, M. (2006): What Do We Know About Competition And Quality In Health Care Markets? NBER working paper No. w12301. National Bureau of Economic Research.
- Godager, G., and Lurås, H. (2009): Dual Job Holding General Practitioners: The Effect of Patient Shortage. Health Economics, forthcoming
- Goldberger, A.S. (1972): Structural Equation Methods in the Social Sciences. Econometrica 40, 979-1001
- Howard, D.H. (2005): Quality and Consumer Choice in Healthcare: Evidence from Kidney Transplantation. *Topics in Economic Analysis & Policy*.
- Hsiao, C. (2003): Analysis of Panel Data, Second edition. Cambridge: Cambridge University Press.
- Jöreskog, K.G. (1977): Structural Equation Models in the Social Sciences: Specification, Estimation and Testing, in P.R. Krishnaiah (ed.): *Applications of statistics*. Amsterdam: North Holland, pp. 265-287.
- Jöreskog, K.G., D. Sörbom, D., du Toit, D. and du Toit. M. (2000): LISREL 8: New Statistical Features. Lincolnwood, IL: Scientific Software International.
- Lantz et al. (1998): Socioeconomic Factors, Health Behaviors, and Mortality. Results From a Nationally Representative Prospective Study of US Adults. The Journal of the American Medical Association 279,1703-1708
- Lee, L.F., and Griffiths, W.E. (1979): The Prior Likelihood and Best Unbiased Prediction in Stochastic Coefficient Linear Models. Working Papers in Econometrics and Applied Statistics No. 1. Department of Economic Statistics, University of New England.
- Luft, H.S. et al. (1990): Does Quality Influence Choice of Hospital? Journal of the American Medical Association 263, 2899-2906.
- Lurås, H., Paulsen, B. and Finnvold, J.E. (2003): De fleste fikk den de ville ha. Samfunnsspeilet 05/2003.
- Ma, C.T.A. (1994): Health Care Payment Systems: Cost and Quality Incentives. *Journal of Economics and Management Strategy* 3, 93-112.
- McFadden, D., 1981: Econometric models of probabilistic choice, in: C.F. Manski and D. McFadden, eds., Structural analysis of discrete data with econometric application, (The MIT Press, Cambridge, MA).
- Papas G., et al. (1993): The Increasing Disparity in Mortality Between Socioeconomic Groups in the United States, 1960 and 1986. The New England Journal of Medicine 329, 103-109.
- Statistics Norway (2005a): Dødelighetstabeller 2004. Oslo: Statistics Norway. http://www.ssb.no/emner/02/02/10/dode/arkiv/tab-2005-04-28-05.html.
- Statistics Norway (2005b): Folkemengde, etter kjønn og ettårig alder. Oslo: Statistics Norway. http://:statbank.ssb.no/statistikkbanken.
- Statistics Norway (2006a): Indeks for levekårsproblemer og delindekser for enkeltindikatorer http://www.ssb.no/vis/emner/03/hjulet/tabell-04.html
- Statistics Norway (2006b): Alders- og kjønnsstandardisert dødelighet. Gjennomsnitt for årene 2001-2005. Dødelighet per 100 000 innbyggere. Konfidensintervall for dødelighetsraten http://www.ssb.no/emner/03/hjulet/tabell-08.1.html
- Train, K. E., (2003): Discrete Choice Methods with Simulation. Cambridge University Press

Table 1: Pattern of Observations

Response pattern	No. of GPs	Freq., %	Cum. freq., %
1111111	441	91.12	91.12
11111	11	2.27	93.39
111111.	11	2.27	95.66
1111	8	1.65	97.31
111	7	1.45	98.76
11	5	1.03	99.79
1111	1	0.21	100.00
	484	100.00	• •

TABLE 2: VARIABLE DEFINITIONS

Variable	Definition, Type of variable	Formula
DEAD	No. of dead persons on GP's list	
EXPDEAD	No. of persons on GP's list expected to die per year	Expected mortality rates based on age distribution of persons on list and population age-specific mortality rates
ACTMORT	Actual no. of mortalities per 1000 persons listed	= DEAD/LISTSIZE
EXPMORT	Expected no. of mortalities per 1000 persons listed	=EXPDEAD/LISTSIZE
EXCMORT	Excess mortality relative to list size	= ACTMORT - EXPMORT
LISTSIZE	GP's actual no. of patients	
DEMAND	No. of persons ranking this GP as most preferred when returning entry form	
DEMAND1	Demand for this GP normalized against GP density in municipality	= DEMAND * GPDENSITY
AGEGP	Age of GP, January 2002	
LEAKRATE	Share of patients switching to another GP.	= no. of persons leaving/LISTSIZE
LOLEAK	$\log(\text{LEAKRATE}/(1-\text{LEAKRATE}))$	
FEMALEGP	Dummy, $=1$ if GP is female	
MARRIEDGP	Dummy, =1 if GP is married	
IMMIGRGP	Dummy, =1 if GP is non-Scandinavian citizen	
SALARY	Dummy, =1 if GP is remunerated by a fixed salary scheme	
SPECGEN	Dummy, =1 if GP is a specialist in general practice	
SPECCOM	Dummy, =1 if GP is a specialist in community medicine	
SPECOTH	Dummy, =1 if GP is a specialist in other kind of medicine	
LEASTCENT	Dummy, =1 if practice in Least central municipality	
LESSCENT	Dummy, =1 if practice in Less central municipality	
CENTRAL	Dummy, =1 if practice in Central municipality	
MOSTCENT	Dummy, =1 if practice in Most central municipality	
PINCOME	Median income (NOK 1000) of persons assigned to this GP in 2001	
PWEALTH	Median wealth (NOK 1000) of persons assigned to this GP in 2001	
PFORMSUB	Share of persons returning forms in 2001 among those assigned to this GP in 2001	
LOSUBMIT	$\log(PFORMSUB/(1-PFORMSUB))$	
PEDUC	Share of persons without finished high-school among those assigned to this GP in 2001	
LOEDUC	log(PEDUC/(1-PEDUC))	
GPDENSITY	No. of GPs per 1000 inhabitans in municipality	

Table 3: Global descriptive statistics

Variable	Obs	Mean	St. Dev	Skewness	Kurtosis	Min	Max
DEAD	3260	5.6254	4.2566	1.2850	5.3638	0	29
EXPDEAD	3275	5.1252	3.2353	1.1384	4.6549	0.0503	23.9037
ACTMORT	3251	4.7629	3.5683	1.9031	11.9554	0	37.9669
EXPMORT	3256	4.2390	2.1982	0.8033	3.3628	0.3781	13.6722
EXCMORT	3251	0.5235	2.6810	2.2873	17.5835	-8.4804	29.4960
DEMAND	3275	825.9289	520.5296	1.1275	4.8392	21	3152
DEMAND1	3275	716.0950	447.6848	1.1181	4.8671	18.3694	2757.1620
LEAKRATE	3251	0.0290	0.0269	6.2202	108.6469	0	0.6660
AGEGP	3275	47.0293	7.5861	0.0610	2.8225	28	70
FEMALEGP	3275	0.3600	0.4801	0.5833	1.3403	0	1
MARRIEDGP	3275	0.6889	0.4630	-0.8159	1.6656	0	1
IMMIGRGP	3275	0.0504	0.2188	4.1111	17.9015	0	1
SALARY	3275	0.0544	0.2267	3.9315	16.4564	0	1
LISTSIZE	3256	1211.0820	401.3513	0.1665	3.4662	123	2687
SPECGEN	3275	0.5597	0.4965	-0.2405	1.0578	0	1
SPECCOM	3275	0.0696	0.2545	3.3821	12.4389	0	1
SPECOTH	3275	0.0345	0.1825	5.1008	27.0180	0	1
LEASTCENT	3275	0.0256	0.1581	6.0012	37.0144	0	1
LESSCENT	3275	0.0116	0.1071	9.1212	84.1960	0	1
CENTRAL	3275	0.0403	0.1967	4.6747	22.8526	0	1
MOSTCENT	3275	0.9224	0.2675	-3.1588	10.9778	0	1
PINCOME	3275	196.5151	23.7784	0.2245	2.7626	130.9460	261.7615
PWEALTH	3275	63.4333	79.2646	1.2259	4.9701	-94.0780	467.3245
PFORMSUB	3275	0.7320	0.2642	-0.7781	2.5387	0.0554	1.0000
PEDUC	3275	0.4446	0.1524	-0.2344	2.4003	0.0213	0.8090

Table 4: Descriptive statistics for the 484 GP-specific means

Variable	Mean	St. Dev	Skew	Kurt	Min	Max	Between variation as share of total, %
DEAD	5.5612	3.6936	1.1570	4.8702	0.0000	22.0816	75.3
EXPDEAD	5.0638	3.2286	1.2459	5.2236	0.3427	22.6007	99.6
ACTMORT	4.7216	2.9549	1.7391	11.2484	0.0000	28.1421	68.6
EXPMORT	4.2065	2.1513	0.8130	3.3556	0.4489	12.4530	95.8
EXCMORT	0.5151	1.8582	3.3390	30.4057	-4.1782	20.3758	48.0
LEAKRATE	0.0308	0.0255	4.4032	36.3194	0.0007	0.2737	90.2
LISTSIZE	1201.6160	401.2103	0.1835	3.4536	152.7857	2620.0000	99.9

Table 5: Descriptive statistics for variables at the municipal level

Municipality	No. of GPs	GPDENSITY
Frogn	7	0.69
Oslo	426	0.87
Stor-Elvdal	2	0.68
Søndre Land	5	0.98
Notodden	10	0.82
Tvedestrand	5	0.84
Vindafjord	4	0.83
Os	12	0.86
Jølster	2	1.01
Ulstein	6	0.91
Overhalla	2	0.55
Beiarn	1	1.57
Porsanger	2	0.69
Total	484	

Table 6: Model A. Mortality equation, GLS estimates

No. of obs.=3251. No. of GPs=484. Obs. per GP.: min=1, mean=6.7, max=7 $\,$

Regressor	Estimate	Std.Err.		
EXPMORT	1.2900	0.0383**		
CENTRAL	-0.6365	0.1656**		
LESSCENT	-0.3897	0.0434**		
LEASTCENT	-0.2359	0.2578		
LOSUBMIT	-0.0496	0.0025**		
LOEDUC	-0.0468	0.0185*		
PINCOME	-0.0097	0.0015**		
PWEALTH	-0.0094	0.0006**		
SPECGEN	0.0156	0.0323		
SPECCOM	0.2164	0.0536**		
SPECOTH	-0.6432	0.0717**		
FEMALEGP	-0.1720	0.0422**		
AGEGP	-0.0547	0.0498		
AGEGPSQ	0.0003	0.0005		
CONST	4.0038	1.0395*		
$\sigma_{\alpha 2}$	1.30	602		
σ_{u2}	2.10	076		
ρ	0.29	940		
R^2				
within	0.0578			
between	0.7104			
overall	0.5071			
Wald chi2(11)	165	868		
p	0.00	000		

^{*)} Significantly $\neq 0$ at the 5 % level

^{**)} Significantly $\neq 0$ at the 1% level

Table 7: Model A: Demand equation: Between GP estimates

No. of obs.=3251. No. of GPs=484. Obs. per GP.: min=1, mean=6.7, max=7 $\,$

Descriptive statistics for Alphahat:

Mean: 0.0000, St.Dev: 1.1546, Skew: 3.5900, Kurt: 37.5611, Min: -3.3982, Max: 13.1143

Demand measure 1: No. of first rankings, unweighted. Demand measure 2: No. of GP density weighted first rankings

	Demand Mea	SURE 1	Demand Mea	SURE 2
Regressor	Estimate	Std.Err.	Estimate	Std.Err.
GPDENS1	-950.3971	445.4716*		
MARRIEDGP	144.8451	48.7506**	124.3429	42.1185**
SPECGEN	158.4471	54.4290**	140.2252	46.9876**
SPECCOM	-156.025	90.4873	-136.4362	78.0651
SPECOTH	-159.9245	125.6506	-134.8392	108.5775
ALPHAHAT	-66.018	19.1214**	-56.6003	16.5214**
IMMIGRGP	-114.8588	104.0845	-95.0004	89.9192
FEMALEGP	-89.3546	48.5764	-75.4430	41.9200
AGEGP	82.8435	29.0487**	72.4525	24.9295**
AGEGPSQ	-0.8065	0.3023**	-0.7066	0.2595**
CONST	-545.2106	824.6051	-1204.6690	586.5191*
R^2				
within	0.0000		0.0000	
between	0.1403		0.1298	
overall	0.1372		0.1265	
	F(10,473) = 7.72		F(9,474) = 7.86	
p	0.0000		0.0000	

^{*)} Significantly $\neq 0$ at the 5% level. **) Significantly $\neq 0$ at the 1% level

Table 8: Model B: Equation for latent demand.

A. Coefficient estimates

Regressor	Estimate	Std.Err.
Quality	1.3583	0.1617**
Health	0.0456	0.0177**
femalegp	-1.0422	0.3881**
immigrant	-0.9545	0.8356

^{**)} Significantly $\neq 0$ at the 1% level

B. VARIANCE AND CORRELATION MATRIX OF LATENT VARIABLES Variances along main diagonal, correlation coefficients below.

	Demand	Quality	Health
Demand	15.6585		
Quality	0.4358	2.4438	
Health	0.0123	-0.4064	460.2335

Appendix A: Model with excess mortality exogenous to demand

The estimation results when excess mortality rates at the level of the GP is entered as an exogenous explanatory variable are presented in Table A.1, panel A based on the dependent variable DEMAND, panel B on the dependent variable DEMAND1. The only difference between the results presented in Table A.1 and the results presented in Table 7 is that ALPHAHAT is replaced by EXCMORT. The latter variable is now assumed to be a valid proxy for quality, and is assumed to be uncorrelated with the error components. We see that the estimated effect of GP density is statistically significant and has the expected sign. We see that being married and being a specialist in general medicine have a positive and statistically significant effect on demand. We also see that AGEGP has a statistically significant effect. The effect of EXCMORT is negative and statistically significant. An interpretation may be that a marginal increase in quality, measured in terms of a reduction in the excess mortality rates, increase the demand facing the GP. The magnitude of the estimated coefficient indicates that a marginal quality increase equivalent to a reduction of the mortality rate of one per thousand, increases the demand facing the GP by 90 persons (Panel A). We note that the estimated coefficients on EXCMORT are larger in absolute value than the coefficients on ALPHAHAT in Table 7. An interpretation of this result, while recalling that the dimension of the coefficients are the same, is that by representing quality by EXCMORT, as in Table A.1, we disregard the variation in health status between GPs. We cannot assess whether high excess mortality is a result of low quality or bad health status of the listed patients.

Table A.1: Model A: Between GP estimates of demand equation. No. of obs.=3251. No. of GPs=484. Obs. per GP.: min=1, mean=6.7, max=7

Demand Measure 1:	No. of first rankings, unweighted.
Demand Measure 2: No.	OF GP DENSITY WEIGHTED FIRST RANKINGS

	Demand meas	SURE 1	Demand measure 2		
Regressor	Estimate	Std.Err.	Estimate	Std.Err.	
GPDENS1	-922.5879	424.9770*			
MARRIEDGP	113.9111	46.7327*	97.3718	40.3584*	
SPECGEN	150.8321	51.9366**	133.7817	44.8165**	
SPECCOM	-156.9416	86.3357	-137.5563	74.4515	
SPECOTH	-182.0114	119.8388	-153.7605	103.5101	
EXCMORT	-90.6113	11.7378**	-78.5167	10.1394**	
IMMIGRGP	-119.6318	99.2157	-99.4600	85.6771	
FEMALEGP	-133.5571	46.6809**	-113.5944	40.2677**	
AGEGP	72.0700	27.7608**	62.9205	23.8148**	
AGEGPSQ	-0.7207	0.2887*	-0.6302	0.2478*	
CONST	-167.7292	788.7412	-851.8743	561.5582*	
R^2					
within	0.0000		0.0000		
between	0.2172		0.2084		
overall	0.1408		0.1323		
	F(10,473) = 13.12		F(9,474) = 13.87		
<i>p</i>	0.0000		0.0000		

^{*)} Significantly $\neq 0$ at the 5% level. **) Significantly $\neq 0$ at the 1% level.

Appendix B: Details on the measurement equations of Model B

The X-measurement model: The left hand side of Figure 2 describes the relations in the X-measurement model. Since latent variables by definition do not have a scale of measurement, the scale of the latent variables are defined by fixing one or more of the factor loadings. In order to make the results from LISREL estimation comparable to the results from Model A we have scaled the ξ variable Quality to be measured in (negative) units of per thousand excess mortality rates. By fixing the factor loading of pincome on Health, we have scaled the ξ variable Health to be measured in units of thousand Norwegian kroner. The interpretation is that the GP's gender

and country of birth is expected to affect the demand directly, without being indicators of either quality or health status of listed patients.

The excess mortality rate from each period is entered as an indicator of GP quality and as indicators of the aggregate measure of health status of listed persons. Other variables included as indicators of quality are dummy variables indicating GP specialization, a dummy variable indicating whether or not the GP is remunerated by means of a fixed salary, a dummy variable indicating whether or not the GP is married and the GP's age. The proportion of listed individuals without finished high school and the proportion of listed individuals who submitted the entry form in 2001 are included as indicators of the aggregate health measure. The estimation results, corresponding to Equation (11), are given in Table B.1.

The estimated factor loading of *Quality* on *married*, *agegp* and *specgen1* are statistically significant, and the interpretation is that married GPs, GPs who are specialists in general medicine is associated with higher quality and that there is a positive relation between GP age and provided quality. All the factor loadings of *Health* are statistically significant, and we see that excess mortality rates are negatively related to *Health*. We see that *Health* is positively related to the share of individuals who submitted the entry form in 2001, and that our latent aggregate health measure is negatively related to the proportion of listed persons with short schooling, as expected.

The estimated covariance matrix of error terms in Equation (11) is given in Table B.2. We have added some restrictions on the correlation of error terms from different x-regressions. With some exceptions error terms from different regressions are uncorrelated. Error terms in regressions on the quality indicators related to GP specialization, specgen1, specoth1 and speccom1, are allowed to be correlated. Further, error terms relating to the excess mortality rates regression from period t are allowed to be correlated with error terms in period t-1.

The Y-measurement model: The right hand side of Figure 2 describes the relations in the Y-measurement model. Here the observable variable demand1 enters the model as an indicator of the η variable Demand. The variable demand1 is identical to the GP density weighted number of first rankings applied in estimation of Model A.

The factor loading of Demand on demand1 is fixed to 100 and the η variable Demand is thus measured in units of hundred first rankings. The log-odds-ratio of the proportion of listed persons leaving the list in period t, $loleak_t$, are entered as indicators of Demand. The estimation results, corresponding to Equation (10), are given in Table B.3. All the factor loadings are statistically significant, and the factor loadings of the $loleak_t$ variables and the factor loading of Demand on demand1 are of opposite sign, as one should expect. The estimated covariance matrix of error terms in Equation (10) is given in Table B.4. No restrictions are imposed on this matrix.

Table B.1: Model B: X-measurement equations
No. of obs.=484.

Regressors	ξ variabli	E: Quality	ξ VARIABLI	E: HEALTH
Dep. var.	Estimate	Std.Err.	Estimate	Std.Err.
excmort1	-1	(Fixed)	-0.0614	0.0127**
excmort3	-1	(Fixed)	-0.0498	0.0121**
excmort4	-1	(Fixed)	-0.0581	0.0123**
excmort5	-1	(Fixed)	-0.0458	0.0120**
excmort6	-1	(Fixed)	-0.0673	0.0127**
excmort7	-1	(Fixed)	-0.0445	0.0119**
loeduc	0	(Fixed)	-0.0283	0.0025**
losubmit	0	(Fixed)	0.0374	0.0158*
pincome	0	(Fixed)	1	(Fixed)
femalegp	0	(Fixed)	0	(Fixed)
immigrgp	0	(Fixed)	0	(Fixed)
married	0.0558	0.0161**	0	(Fixed)
agegp	1.0471	0.2692**	0	(Fixed)
salary1	0.0033	0.0079	0	(Fixed)
specgen1	0.0658	0.0174**	0	(Fixed)
specoth1	-0.0056	0.0064	0	(Fixed)
speccom1	0.0056	0.0086	0	(Fixed)

^{*)} Significantly $\neq 0$ at the 5% level. **) Significantly $\neq 0$ at the 1% level.

Table B.2: Model B: Error covariance matrix of X-measurement equations

excmort1 0.7120 excmort3 0.1371 0.6403 excmort4 - 0.0727 0.6520 excmort5 - - -0.0077 0.6397 excmort6 - - - -0.1060 0.6625							
excmort3 0.1371 0.6403 excmort4 - 0.0727 0.6520 excmort5 - - -0.0077 0.6397 excmort6 - - - - -0.1060 0.6625		excmort1	excmort3	excmort4	excmort5	excmort6	excmort7
excmort4 - 0.0727 0.6520 excmort5 - - -0.0077 0.6397 excmort6 - - - -0.1060 0.6625	excmort1	0.7120					
excmort50.0077 0.6397 excmort60.1060 0.6625	excmort3	0.1371	0.6403				
excmort60.1060 0.6625	excmort4	_	0.0727	0.6520			
	excmort5	_	_	-0.0077	0.6397		
excmort7 0.0046 0.6026	excmort6	_	_	_	-0.1060	0.6625	
	excmort7	_	_	_	_	0.0046	0.6026

	loeduc	pincome	losubmit	salary1	married	agegp	specgen1	specoth1	speccom1
loeduc	0.2931								
pincome	-	0.1951							
losubmit	_	_	0.9866						
salary1	_	_	_	0.9995					
married	_	_	_	_	0.9648				
agegp	_	_	_	_	_	0.9550			
specgen1	_	_	_	_	_	_	0.9575		
specoth1	_	_	_	_	_	_	-0.0043	0.9979	
speccom1	_	_	_	_	_	_	0.1463	-0.0082	0.9988

TABLE B.3: MODEL B: Y-MEASUREMENT EQUATIONS

No. of obs.=484.

	η variable: Demand					
Dep. var.	Estimate	Std.Err.				
loleak1	-0.1671	0.0165**				
loleak3	-0.1542	0.0167**				
loleak4	-0.2337	0.0759**				
loleak5	-0.1600	0.0465**				
loleak6	-0.2541	0.0899*				
loleak7	-0.0850	0.0494				
demand1	100	(Fixed)				

^{*)} Significantly $\neq 0$ at the 5 % level.

Table B.4: Model B: Error covariance matrix of Y-measurement equations

	loleak1	loleak3	loleak4	loleak5	loleak6	loleak7	demand1
loleak1	0.4024						
loleak3	0.2469	0.4318					
loleak4	0.1819	0.2454	0.9014				
loleak5	0.2058	0.2489	0.5517	0.8818			
loleak6	0.1312	0.1983	0.3027	0.4874	0.9123		
loleak7	0.2182	0.2768	0.1667	0.2079	0.1671	0.9673	
demand1	-0.0235	0.0049	-0.0101	-0.0394	0.0359	-0.0858	0.2182

Appendix C: Alternative LISREL model. Model C

^{**)} Significantly $\neq 0$ at the 1% level.

Table C.1: Model C: Equations for latent dependent variables.

A. Coefficient estimates

		Regressors								
	Qu.	ALITY	IMMIC	GRANT	MA	RRIED	AG	EGP	FEMA	ALEGP
Dep. var.	Est.	Std.Err.	Est.	Std.Err.	Est.	Std.Err.	Est.	Std.Err.	Est.	Std.Err.
Demand	0.9759	0.1487**	-1.2020	0.7935	1.2539	0.3748**	0.0293	0.0226	-0.6711	0.3733
Health	-0.0019	0.0019	0	(Fixed)	0	(Fixed)	0	(Fixed)	0	(Fixed)

^{*)} Significantly $\neq 0$ at the 5% level. **) Significantly $\neq 0$ at the 1% level.

 $\begin{array}{c} {\rm B.\ VARIANCE\ AND\ CORRELATION\ MATRIX\ OF\ LATENT\ VARIABLES} \\ {\rm Variances\ along\ main\ diagonal,\ correlation\ coefficients\ below} \end{array}$

	Quality	Demand	Health
Quality	2.1382		
Demand	0.3426	17.3484	
Health	-0.1268	-0.0434	0.0005

Table C.2: Model C: Y-measurement equation

No. of obs.=484.

	Quality		Demand		Health	
DEP VAR	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.
demand1	0	(Fixed)	100.0000	(Fixed)	0	(Fixed)
loleak1	0	(Fixed)	-0.1576	(0.0178)**	0	(Fixed)
loleak3	0	(Fixed)	-0.1650	(0.0186)**	0	(Fixed)
loleak4	0	(Fixed)	-0.3930	(0.0598)**	0	(Fixed)
loleak5	0	(Fixed)	-0.2673	(0.0392)**	0	(Fixed)
loleak6	0	(Fixed)	-0.4151	(0.0663)**	0	(Fixed)
loleak7	0	(Fixed)	-0.2038	(0.0339)**	0	(Fixed)
speccom1	0.0016	(0.0092)	0	(Fixed)	0	(Fixed)
specoth1	-0.0046	(0.0069)	0	(Fixed)	0	(Fixed)
specgen1	0.0452	(0.0181)**	0	(Fixed)	0	(Fixed)
salary1	0.0055	(0.0085)	0	(Fixed)	0	(Fixed)
excmort1	-1.0000	(Fixed)	0	(Fixed)	-39.3361	(9.9107)**
excmort3	-1.0000	(Fixed)	0	(Fixed)	-28.0050	(9.1767)**
excmort4	-1.0000	(Fixed)	0	(Fixed)	-35.8985	(9.4614)**
excmort5	-1.0000	(Fixed)	0	(Fixed)	-24.0063	(9.1388)**
excmort6	-1.0000	(Fixed)	0	(Fixed)	-44.6527	(9.8519)**
excmort7	-1.0000	(Fixed)	0	(Fixed)	-23.3650	(9.0128)**
pincome	0	(Fixed)	0	(Fixed)	1000.0000	(Fixed)
pwealth	0	(Fixed)	0	(Fixed)	439.5965	(174.5656)**
loeduc	0	(Fixed)	0	(Fixed)	-26.9137	(2.5591)**
losubmit	0	(Fixed)	0	(Fixed)	37.1699	(15.3537)**

^{*)} Significantly $\neq 0$ at the 5% level. **) Significantly $\neq 0$ at the 1% level.

Table C.3: Model C: Error covariance matrix in Y-measurement equations

	demand1	loleak1	loleak3	loleak4	loleak5	loleak6	loleak7
demand1	21999.3356						
loleak1	9.7168	0.2958					
loleak3	51.4310	0.1182	0.1773				
loleak4	283.0998	_	0.0333	5.9740			
loleak5	174.9324	_	_	1.5528	2.0378		
loleak6	339.9960	_	_	_	1.4497	8.5124	
loleak7	153.7280	_	_	_	_	-0.1123	2.7219

	excmort1	excmort3	excmort4	excmort5	excmort6	excmort7
excmort1	6.2133					
excmort3	1.0092	3.9033				
excmort4	_	0.3908	4.4066			
excmort5	_	_	-0.1208	3.8241		
excmort6	_	_	_	-0.8207	5.1367	
excmort7	_	_	_	_	-0.0872	3.2198

	speccom1	specoth1	specgen1	salary1	pincome	pwealth	loeduc	losubmit
speccom1	0.0635							
specoth1	-0.0005	0.0358						
specgen1	0.0191	-0.0009	0.2444					
salary1	_	_	_	0.0544				
pincome	_	_		_	87.6548			
pwealth	_	_	_	_	_	6131.2983		
loeduc	_	_	_	_	_	_	0.1696	
losubmit	_	_	_	_	_	_	_	47.5383

Dual Job Holding General Practitioners: The Effect of Patient Shortage*

Geir Godager¹** and Hilde Lurås²

Summary

In 2001, a listpatient system with capitation payment was introduced in Norwegian general practice. After an allocation process where each inhabitant was listed with a general practitioner (GP), a considerable share of the GPs got fewer persons listed than they would have preferred. We examine whether GPs who experience a shortage of patients to a larger extent than other GPs seek to hold a second job in the community health service even though the wage rate is low compared to the wage rate in general practice. Assuming utility maximization, we model the effect of patient shortage on a GP's decision to contract for a second job in the community health service. The model predicts a positive relationship between patient shortage and participation in the community health service. This prediction is tested by means of censored regression analyses, taking account of labour supply as a censored variable. We find a significant effect of patient shortage on the number of hours the GPs supply to community health service. The estimated marginal effect is 1.72 hours per week.

Keywords: General practice, Patient shortage, Dual job holding.

¹Institute of Health Management and Health Economics, University of Oslo.

² Helse Sør-Øst Health Services Research Centre, Akershus University Hospital.

^{*} This paper was accepted for publication in *Health Economics* on August 19th 2008.

^{**} Corresponding author: Address: P.O. Box 1089 Blindern, N-0317 Oslo, Norway. E-mail address: g.c.godager@medisin.uio.no . Tel.:+47 22 84 50 29; fax: +47 22 84 50 91

1. Introduction and background

Physicians holding dual or multiple jobs can be observed in both developing and high-income countries. In various institutional settings, incentives for dual job holding may emerge as a result of complementarities between jobs. Opportunities for cream skimming or self-referrals of patients from the public sector to own private practice seem to encourage dual practice (Gonzalez, 2004, Barros and Olivella, 2005, Iversen, 1997, Eggleston and Bir, 2006). While dual practice is strictly regulated or prohibited in some countries such as Canada, it is often a result of tight public health care budgets in low income countries (Bir and Eggleston, 2003, Jan et al. 2005).

A decision to hold a second job may also be driven by constraints in the physician's primary job. Shishko and Rostker (1976) showed that restrictions on hours worked in the primary job imply incentives to hold a second job (moonlighting), and that individuals experiencing such restrictions may accept wage rates that are lower than the wage rate in the primary job. In a study of moonlighting decisions by resident physicians, Culler and Bazzoli (1985) found that the number of hours employed in resident programs had a negative effect on the supply of moonlighting labour. In this paper we explore this further by questioning whether general practitioners (GPs) experiencing patient shortage in their private practice to a larger extent than unconstrained GPs seek to hold a second job in community health service, even though the hourly wage rate in this job is considerably lower than the hourly income from working in general practice. By means of an empirical analysis based on data from 387 GPs practicing in 26 municipalities in Norway and two districts in Oslo (the capital city), we find that GPs experiencing patient shortage supply more hours to the community health service compared to GPs who are not experiencing a shortage of patients in their private practice.

The paper proceeds as follows. In Section 2 we describe the organization of primary health care in Norway and in section 3 we model the effect of patient shortage on a GP's decision to contract for a second job in the community health service. Section 4 describes the data, while empirical specification and estimation methods are presented in section 5.

The results are given in Section 6, while section 7 concludes and discusses policy implications of the findings.

2. The organization of primary health care in Norway

Norway has a national health service which is mainly tax-financed. The organization of primary health care is the responsibility of the municipalities, which is the lowest governmental level. Due to the implementation of a listpatient system (the regular general practitioner scheme) in 2001, every inhabitant is registered with a GP. 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 from the National Insurance Administration and from patients' co-payments. The fee for service component is paid according to a fixed fee schedule negotiated between the state and the Norwegian Medical Association. The fees depend on the duration of the consultation, whether certain types of examinations and laboratory tests are initiated during the consultation, and whether the GP is a specialist in general medicine. The patients pay a fixed fee per consultation (amounting to € 14.70 in 2002) with an annual ceiling.

Privately practicing GPs are responsible for providing general medical services to persons listed in their practices. However, certain segments of the primary care sector are not part of the domain of general practice:

- Administration of the primary health care sector
- Public health services such as the preparation of infectious disease plans
- Certain medical tasks such as routine examinations of infants and children at childcare centers and schools, and regular medical care at nursing homes, prisons and other institutions

These tasks are part of the community health service, and carried out by GPs working in part-time positions in the municipality. Positions in community health service are remunerated according to a fixed salary scheme negotiated in centralized wage settlements between the Norwegian Medical Association and the Norwegian Association of Local and Regional Authorities. These positions have long been regarded as unprofitable among the GPs. The hourly wage rate for GPs working in community health

service is estimated to be between 38% and 66% of the wage rate in private practice (Godager and Lurås, 2005). According to prevailing regulations, the municipalities can require GPs to work up to 7.5 hours per week with community health service tasks. This requirement has important implications for our empirical analysis, and will be discussed in more detail in Sections 3 to 5.

3. A model for the effect of patient shortage on a GP's decision to contract for a second job in community health service

We model the effect of patient shortage on a representative GP's decision to contract for a second job in community health service. We assume that all persons listed with the GP are homogeneous with respect to their need for health care, and that the GP provides a fixed amount of a composite health service to every person on the list. The GP faces a market demand, n^D , of persons who request to be listed in the practice. We further assume that the GP is not able to affect n^D , and that the market demand constitutes an exogenous upper limit to the number of persons listed. A GP experiences patient shortage when the market demand, n^D , is less than the GP's preferred list size, n^D .

In the model we take account of the municipalities' option to require GPs to work up to 7.5 hours per week in community health service by including the restriction that the GP cannot work less than Φ hours per week with these tasks, where Φ ranges from zero to 7.5 hours. The upper limit of 7.5 hours is part of central regulations. However, the local health authorities may set Φ lower than 7.5 hours.

The physician maximizes a quasi linear objective function U=c+v(l), where c is consumption and l is leisure. We assume v'(l)>0, v''(l)<0 and $\lim_{l\to 0}v'(l)=\infty$.

We disregard savings, and the amount of consumption equals the GP's monetary income. The GP's total income from working in private practice (the primary job) and in community health service (the second job) is $qn + w\tau$, where n is the number of persons listed and q is the income per listed person²; w is the fixed hourly wage rate the GP

receives if contracting for a second job in the community health service; and τ is the number of hours worked in this job. Leisure is defined by $l = T - nt - \tau$, where T is the total time endowment and t is the strictly positive time input per person listed in the GP's practice.

One may think of the model as dynamic with sequential decision making, where the GP first decides on the desired number of patients, n^P , conditional on minimum hours in the community health service, $\Phi = \tau$. In the next step, demand, n^D , is revealed and determines whether patient shortage occur. In the last step, the GP decides his or her optimal number of hours in the community health service. A static formulation of the model may be specified without loss of generality, and we formulate the maximization problem as one where the GP choose optimal n and τ given that the number of persons requesting to be listed in the GP's practice (n^D) and the number of hours the GP is required to work in the community health service (Φ) restrict the opportunity set. The decision problem can be expressed as:

$$Max_{n,\tau} qn + w\tau + v(T - nt - \tau)$$

s.t.

- (i) $0 < n \le n^D$
- (ii) $\Phi \leq \tau$

The problem is solved by means of concave programming (details can be found in the appendix). Assuming that the wage rate in the community health service is lower than the income per time unit in private practice, $w < \frac{q}{t}$, we can make the following predictions:

Unconstrained GPs have a higher marginal utility of leisure than constrained GPs.
 Hence, the GPs without patient shortage do not contract for more than the required number of hours (Φ) with the community health service. The unconstrained GPs

- adjust the number of listed persons such that the marginal income equals the marginal utility of leisure.
- Constrained GPs of type 1 are defined as those who have a lower marginal utility of leisure than GPs without patient shortage, and a higher marginal utility of leisure than the wage rate offered by the community health service. These GPs respond to patient shortage by consuming more leisure and will not contract for more than the minimum requirement (Φ).
- Constrained GPs of type 2 are defined as those who have a lower marginal utility of leisure than the wage rate offered by the community health service at $\tau = \Phi$. These GPs will obtain a higher utility if they contract for more than Φ hours with the community health service. Their optimal number of hours contracted for are determined such that the marginal utility of leisure equals the wage rate in the community health service. The optimal number of hours contracted for is a strictly declining function of the market demand for being listed with the GP.

The predictions are summarized in Table I.

Based on the predictions from the theoretical model, the following hypotheses will be tested in the empirical section:

- 1: Constrained GPs contract for more hours with the community health service than unconstrained GPs.
- **2:** Constrained GPs with short lists contract for more hours with the community health service than constrained GPs with long lists.

Table I: Effects of patient shortage on the number of hours contracted with the community health service (*denotes solution values).

Regime		v'(l*)	τ*	$\frac{d\tau^*}{dn^D}$
Unconstrained: n^P	$= n^* < n^D$	$w < v'(T - n * t - \Phi) = \frac{q}{t}$	$ au^* = \Phi$	$\frac{d\tau^*}{dn^D} = 0$
Constrained:	type 1	$w < v'(T - n^D t - \Phi) < \frac{q}{t}$	$ au^* = \Phi$	$\frac{d\tau^*}{dn^D} = 0$
$n^* = n^D$	type 2			$\frac{d\tau^*}{dn^D} = -t < 0$

4. Description of the data

Because the municipalities' organization of primary health care differ according to the centrality of the municipality, we randomly selected the included municipalities according to a centrality measure developed by Statistics Norway³ (Statistics Norway, 1994). This sampling design was chosen in order to include a representative sample of Norwegian municipalities in the analysis.

In every selected municipality we asked the health administration to fill in a postal questionnaire on the number of hours each GP had contracted for work in the community health service and on their wage rate when performing these tasks. This information was merged with registered information on the GPs and the GPs' patient lists from The National Insurance Administration. The response rate on the survey was 100 %. The final sample includes 387 self-employed GPs from 26 municipalities and two districts in Oslo.

The number of hours per week the GPs work in the community health service (WEEKHRS) ranges from zero to 22 (Figure I). Nearly 85 % of the GPs contract for 7.5 hours or less and only 10 % do not have part-time positions in the community health service. In 14 of the municipalities every GP participate in community health service tasks. A majority of the GPs (90 %) receive an hourly wage rate (WAGERATE) in the

range of \in 23 to \in 27. Because the wage rate is negotiated in central wage settlements and the range of the wage rates is quite small, we find it reasonable to assume it to be exogenous to the individual GP. For 80 of the GPs, information on the wage rate was missing, so the mean wage rate among the other GPs in the municipality was applied.

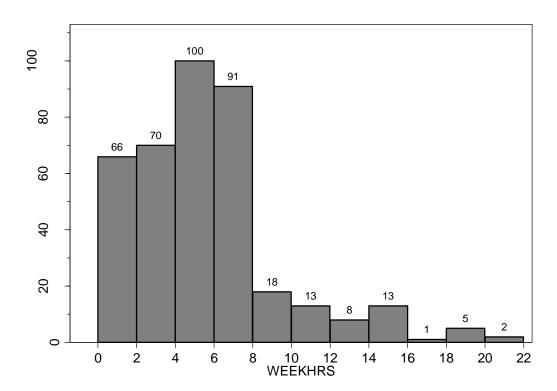


Figure I. Histogram of weekhours contracted with the community health service

Prior to the 2001 reform, the GPs were asked to specify the number of persons they preferred to have on their lists. The GPs are allowed to adjust their preferred list size continuously, and this individual list size limit is public information. Preferred list size is updated in the database in January and July every year. To measure whether a GP experiences patient shortage or not, we compute an indicator (SHORTAGE) which equals 1 if the GP failed to achieve the preferred list size in both January and July 2002 (Iversen and Lurås, 2002). The survey data describes the contracts prevailing in June 2002. The indicator of patient shortage is therefore assumed to be predetermined and exogenous, in the empirical specification.

From the theoretical model, we predicted a relationship between list size and the number of hours the constrained GPs have contracted for work in the community health service. However, since the GPs not experiencing patient shortage can increase their list size, and it follows from the model that increasing the list size is preferable to contracting for work in the community health service, the number of hours supplied by unconstrained GPs is not expected to depend on list size. We include an interaction term (LISTSHORT) consisting of the GP's list size multiplied by the shortage variable to account for this relationship.

To account for observable heterogeneity, we introduce variables describing characteristics of the GP and the GP's patient list. These variables can be interpreted as taste shifters in the utility function. We include four variables characterizing the GP: the GP's gender, the GP's age, and whether the GP has achieved specialist status in general medicine or in community medicine.

From the literature we know that female physicians carry out longer consultations than male GPs do (Iversen and Lurås, 2000, Kristiansen and Mooney, 1993, Langwell 1982). We also know that females, to a larger extent than males, prefer to work part-time (Paulsen and Fjermestad, 1996). Females and males may therefore have different views regarding tasks in community health service. We account for the GP's gender by including a dummy variable FEMALE that equals 1 if the GP is female. To account for differences in medical experiences among GPs, we include the GP's age (GPAGE). GPs may specialize in general medicine or in community medicine. Specialists in general medicine receive a higher fee per consultation from the National Insurance than GPs without this speciality. It is therefore plausible to assume that they have a higher reservation-wage and hence, it is less likely that specialists in general medicine contract for more than the required amount of hours. To indicate the GP's specialisation we include two dummy variables: SPECGEN equals 1 if the GP is a specialist in general medicine and SPECCOM equals 1 if the GP is a specialist in community medicine. Specialists in community medicine are expected to contract for more hours with the community health service than the other GPs.

The observed heterogeneity of the GP's patient list is account for by including the proportion of females and the proportion of old persons on the list. From the literature we know that females and elderly patients on average visit their GPs more often than men and younger patients do (Elstad, 1991). The proportion of female persons on the list (PROPFEM), and the proportion of elderly (PROPOLD) take account of the patient load on a GP's list. Because GPs with a heavier than average patient load may work longer hours, GPs with many females and many elderly on their lists are expected to contract for fewer hours in the community health service. Descriptive statistics of the whole sample, of GPs contracting for less or equal to 7.5 hours, and for GPs contracting for more than 7.5 hours, are given in Table II.

Table II: Descriptive statistics.

	Definition	WEEKHRS<=7.5 n=327	WEEKHRS>7.5 n=60	All GPs	N=387	
Cor	ntinuous variables	Mean (St.dev)	Mean (St.dev)	Mean (St.dev)	Min	Max
WEEKHRS	Hours per week contracting with the c.h.s.*	4.07 (2.36)	12.15 (3.61)	5.32 (3.90)	0	22
WAGERATE	Hourly wage rate (€) in the c.h.s.*	24.74 (1.49)	25.59 (4.19)	24.87 (2.15)	15.25	40.66
PREFLIST	Preferred list size	1447 (376)	1257 (366)	1418 (380)	300	2500
LISTSIZE	Actual list size	1355 (394)	1160 (351)	1325 (393)	197	2798
LISTSHORT	Interaction term of LISTSIZE and SHORTAGE	820.81 (662.75)	764.38 (590.41)	812.06 (651.66)	0	2082
PROPFEM	Proportion of females on list	0.51 (0.10)	0.50 (0.10)	0.51 (0.10)	0.25	0.86
PROPOLD	Proportion of elderly on list	0.11 (0.06)	0.12 (0.07)	0.11 (0.06)	0.00	0.35
GPAGE	GP age (January 2002)	46.42 (9.22)	47.40 (7.21)	46.57 (8.93)	27	69
D	ummy variables	Mean (Count)	Mean (Count)	Mean (Count)		
SHORTAGE	= 1 if constrained GP (PREFLIST>LISTSIZE)	0.65 (214)	0.70 (42)	0.66 (256)		
SPECCOM	= 1 if specialist in community medicine	0.04 (13)	0.18 (11)	0.06 (24)		
SPECGEN	= 1 if specialist in general medicine	0.59 (192)	0.60 (36)	0.59 (228)		
FEMALE	= 1 if GP is female	0.27 (89)	0.20 (12)	0.26 (101)		

^{*} Community health service

5. Empirical specification and estimation

From the theoretical model in Section 3 we predicted that, ceteris paribus, a GP who experiences patient shortage contracts for more hours in the community health service than GPs not experiencing patient shortage, and that constrained GPs contract for more hours the shorter their lists are. In Section 4 we formulated some hypotheses on the effect of exogenous variables on the number of hours GPs contract for with the community health service. These hypotheses will be tested by means of estimation on reduced form equations.

In standard labour markets, individuals who choose not to offer their working capacity to the market will be observed with a labour supply of zero hours, i.e., the variable is censored in zero. Because the municipalities can require GPs to contract for up to 7.5 hours per week we must allow the GP's specific censoring threshold to vary in the interval [0,7.5] hours. When the dependent variable is censored, a linear regression analysis will not give consistent estimates (Tobin, 1958).

Demographic and cultural characteristics and the way primary care is organised are likely to influence a GP's practice style and hence our observations of GPs at the municipality level. An appropriate choice of empirical model and estimation method should therefore take account of unobserved municipality-specific heterogeneity. Since we have repeated observations from the same municipality, we are able to take account of such heterogeneity by applying fixed or random effects estimation. We estimate two censored regression models, one without heterogeneity and one with municipality fixed effects⁴. The estimation is performed using the interval regression module in the computer software STATA 9.0. It should be noted that fixed effect tobit estimation does not yield consistent estimators when the number of municipalities goes to infinity and the number of GPs is finite. On the other hand, consistency is ensured with a finite number of municipalities when the number of GPs goes to infinity (Hsiao, 2003, p 243). Because the number of municipalities is much smaller than the number of GPs, we argue that the application of a fixed effects tobit model is well-founded.

Let \tilde{y}_{ij} denote the latent number of hours GP i practicing in municipality j prefers to work in the community health service. We may interpret \tilde{y}_{ij} as the optimal labour supply resulting from utility maximization without restricting labour supply to be positive (Amemiya, 1984, Maddala, 1983). Further, let y_{ij} denote the number of hours contracted with GP i in municipality j. When $y_{ij} \in \{0,7.5\}$, we do not know whether the contract is a result of the GP's own choice or if the GP is required to contract with the municipality. If a GP has contracted for y_{ij}^0 hours, $y_{ij}^0 \in [0,7.5]$, the only information we have is that $\tilde{y}_{ij} \leq y_{ij}^0$. For $y_{ij} > 7.5$, however, the contract is assumed to be a result of the GP's own decision, and observations of $y_{ij} > 7.5$ therefore enter the regression equation as a continuous variable.

We then have: $y_{ij} = \tilde{y}_{ij}$ if $\tilde{y}_{ij} > 7.5$ and $y_{ij} = y_{ij}^0$ if $\tilde{y}_{ij} \le 7.5$ where y_{ij}^0 is the individual specific censoring threshold. Note that compared to an approach where every observation is assigned a censoring threshold equal to 7.5 hours, assigning an individual specific threshold implies a more efficient use of information on the latent labour supply. We therefore assign the individual specific thresholds according to the following rule:

$$y_{ij}^0 = y_{ij}$$
 if $\tilde{y}_{ij} \le 7.5$
 $y_{ij}^0 = 0$ otherwise

We may now specify our censored regression model:

$$y_{ij} = \beta' x_{ij} + \varepsilon_{ij}$$
 if $RHS > 7.5$
 $y_{ij} = y_{ij}^{0}$ otherwise

where β is a vector of parameters and x_{ij} is a vector of explanatory variables, including municipality dummies, assumed to be uncorrelated with the residuals, ε_{ij} . The residuals are assumed to be normally distributed with a zero mean and a common variance σ^2 .

6. Results

The results from the tobit regression analyses are reported in Table III. By means of a likelihood ratio test we find that the model without heterogeneity is rejected in favour of the fixed effects model. In the following we therefore discuss the results from the fixed effects model. The results indicate that the constrained GPs contract for more hours with the community health service, and that the longer the constrained GPs' lists are, the fewer hours are contracted for. The results are consistent with the predictions from our theoretical model in Section 3. The effect of wage rate is statistically significant and the coefficient has the expected sign. At first glance the marginal effect may seem large: increasing the hourly wage rate by a marginal EURO increases the preferred number of working hours by one hour and ten minutes per week. However, this effect is the predicted increase in the latent labour supply, and not an effect on hours actually supplied. Owing to the large number of censored observations, the effect on the number of hours actually supplied will be considerably lower. In the right column of the table we therefore present the marginal effects of the explanatory variables on the number of hours actually supplied, conditional on the number of hours being more than 7.5. We see that, conditional on working more than 7.5 hours (WEEKHRS>7.5), increasing the wage rate by one EURO increases the GPs' observed labour supply by 0.154 hours per week. This result corresponds to a conditional supply elasticity with respect to wage rate of 0.32. When standard errors are calculated by means of the delta method⁵, the 95 % confidence interval of the conditional supply elasticity is [0.018 - 0.629], which is comparable in magnitude to physician labour supply elasticities found in other studies (Rizzo and Blumenthal, 1994, Showalter and Thurston, 1997, Sæther 2005, Baltagi et al. 2005).

Table III Results from censored regression analysis

The estimated effect of explanatory variables on the number of hours supplied to the community health service (WEEKHRS). Robust Standard errors in parenthesis. N=387. Observations per municipality: min=1, mean=14.3, max=108

Regression model	No heter	rogeneity	Municipality	fixed effects		
	Tobit coefficients	Marginal effect†	Tobit coefficients	Marginal effect†		
SPECCOM	11.279	2.351	7.694	1.220		
	(3.209)**	(0.830)**	(3.064)*	(0.570)*		
SPECGEN	1.050	0.172	3.652	0.476		
	(2.221)	(0.363)	(2.057)	(0.264)		
SHORTAGE	11.857	1.823	14.217	1.720		
	(4.605)**	(0.662)**	(4.526)**	(0.510)**		
LISTSHORT	-0.009	-0.001	-0.010	-0.001		
	(0.003)**	(0.001)**	(0.003)**	(0.000)**		
FEMALE	-1.932	-0.312	-0.709	-0.093		
	(3.391)	(0.536)	(3.232)	(0.420)		
GPAGE	-0.121	-0.020	-0.121	-0.016		
	(0.131)	(0.022)	(0.123)	(0.016)		
PROPOLD	25.887	4.269	6.634	0.877		
	(19.493)	(3.201)	(17.028)	(2.245)		
PROPFEM	1.471	0.243	-1.392	-0.184		
	(14.583)	(2.404)	(13.310)	(1.759)		
WAGERATE	0.386	0.064	1.165	0.154		
	(0.517)	(0.085)	(0.560)*	(0.074)*		
CONSTANT	-18.558 (14.901)		-36.807 (16.685)*			
Log pseudo-likelihood	-334	.296	-308.844			
P-value	0.0	000	0.000			
Likelihood ratio test of model without heterogeneity		LR chi2 (26) = 50.90 P-value= 0.002				
Censoring summary	327 left-censored observations					
	60 uncensored observations					

^{*} The estimated parameter is significantly different from zero at the 5% level in a two-tailed test

^{**} The estimated parameter is significantly different from zero at the 1% level in a two-tailed test

 $[\]dot{\tau} \frac{\delta}{\delta x} E(y|y>7.5)$ For the dummy variables, marginal effects are for a discrete change of dummy variable from 0 to 1

The marginal effect of patient shortage indicates that the effect of experiencing patient shortage on the conditional supply of hours is 1.72 hours per week. The marginal effect of the interaction term (LISTSHORT) indicates that increasing the constrained GP's list size by 100 persons reduces the conditional supply of hours by 0.1 hours per week. We also observe a significant effect from being a specialist in community medicine (SPESCOM). As expected, specialists in community medicine supply more hours than GPs without this speciality, and the estimated marginal effect is 1.22 more hours per week.

Robustness checks

The use of tobit models to estimate labour supply functions has been criticised, first, because of the assumption that the observed labour supply is truncated normal and second, for not separating the decision of whether to participate in the labour market from the decision of how many hours to supply (Cragg, 1971, Melenberg and Van Soest, 1996). In situations where labour supply is clustered around 35 to 40 hours per week, two-part models with less restrictive assumptions where the model for working hours among individuals with positive supply is separated from the model of whether to participate are shown to be more appropriate (Moffit, 1999, Mroz, 1987). This type of clustering distant from the censoring threshold does not appear in our data. However, as a robustness check we estimated a two-step Heckman error correction model. In the first step we estimated the probability of working more than 7.5 hours per week by means of a probit model, and in the second step, including the inverse mills ratio as an explanatory variable, we estimated the supply of hours conditional on working more than 7.5 hours. In this analysis the effect of WAGERATE and SPECCOM was found to have a significant effect only on the probability of working more than 7.5 hours per week, whereas, SHORTAGE, and LISTSHORT had significant effects on both the probability of working more than 7.5 hours per week and on the supply of hours conditional of working more than 7.5 hours per week.

We also estimated the probability of working more than 7.5 hours by means of fixed effects logit estimation. One advantage with the logit model is that conditional fixed effect estimation yields estimators that are consistent when both the number of municipalities and the number of GPs goes to infinity. The logit estimates confirm the results from the fixed effect tobit estimation.

To check whether our results are sensitive to the choice of shortage indicator, we have estimated several models with different definitions of patient shortage including an indicator equal to 1 if the GP's list was shorter than preferred in January 2002 only, and in July 2002 only. The results from these analyses are similar to the results presented in Table III. We also applied a shortage indicator equal to 1 if the discrepancy between preferred and actual list size in July 2002 was more than ten people. The effect of patient shortage is slightly strengthened in these analyses, and the p-value is somewhat lower.⁶

7. Conclusion and policy implications

Results from our empirical analysis indicate that GPs experiencing patient shortage in their private practices contract for more hours with the community health service, and that the shorter their patient lists are, the more working hours they supply. If we interpret contracting for more than the minimum number of hours required by the municipality as holding a second job, our results are in accordance with the literature on multiple job holding (Shishko and Rostker, 1976, Culler and Bazzoli, 1985). Earlier studies on Norwegian GPs have revealed that there is a positive relationship between patient shortage and service provision (Iversen and Lurås, 2000, Iversen, 2004). In the current paper we find that GPs experiencing patient shortage contract for more hours of community health service than unconstrained GPs. Hence, it seems that GPs compensate for a shortage of patients in several ways.

The fact that we make no attempt to model the demand side for labour in community health service might be considered to be a limitation of this study, and including the demand side is an idea for further research. It is well known, however, that for many years Norwegian municipalities have had problems attracting qualified physicians to positions in community health service. The main reason is that these positions are regarded as less profitable among GPs, reflecting that the hourly wage rate is considerably lower than the income from one hour in private practice. Our results indicate both that the wage rates affect supply and that GPs will work in the community health service at the going wage rate if they are experiencing patient shortage in their private practice. One important policy implication of this finding is that for a given level of wage rates, recruiting GPs to positions in the community health service will be easier if a considerable share of the GPs experience patient shortage.

One may argue that low wage rates in community health service may discourage equal distribution of health care services: People living in areas where a large share of the GPs experience patient shortage will have better access both to general practice and to

services provided by the community health service compared to people living in areas where few GPs experience patient shortage.

The analysis also shows that a listpatient system with capitation payment that was implemented to encourage continuity of care in general practice may have an unexpected influence on GPs' involvement in community health service. Hence, it seems that the organization of general practice may cause spill-over effects in another part of the health care system. This result is in accordance with the literature on spill-over effects, showing that the incentive structure in one market may have implications for behaviour in other markets (Yip, 1998, Baker et al. 2004).

Acknowledgements

Thanks to Eline Aas, Erik Biørn, Sverre Grepperud, Tor Iversen, Erik Magnus Sæther and two anonymous reviewers for suggestions that have contributed to improvements of the paper. We are grateful to participants at The 25th Meeting of the Nordic Health Economists' Study Group, Reykjavik, August 2004 and The 5th European Conference on Health Economics in London, September 2004, 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 authors alone are 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 authors have no conflicts of interests with regard to this research.

¹ This may be an unreasonable assumption in the long run. In the short run, however, it seems likely that GPs do not have discretionary power to change the demand.

² Because we assume constant service intensity, the income from private practice is proportional to the number of listed persons n.

³ The centrality indicator captures the size of the population, population density, and the distance to the nearest city of a certain size.

⁴ We also estimated a random effects model. STATA uses adaptive Gauss-Hermite quadrature to compute the log-likelihood and the derivatives when estimating random effects tobit regressions. We do not present the results from this regression since the estimated coefficients were sensitive to the number of quadrature points applied, indicating that the quadrature approximation is not accurate.

⁵ The electric conditions that the quadrature approximation is not accurate.

The elasticity and its standard error are provided by the software.

⁶ We also tested other indicators of shortage. When the discrepancy between preferred and actual list size is increased to more than 35 people, the effect of patient shortage is not significantly different from zero. The reason is probably that the number of GPs classified as experiencing patient shortage is reduced when the patient shortage criterion is strengthened, which implies increasing standard errors in the estimated parameters. Most likely this happens because of the small sample size.

Appendix: Derivation of the results in Section 2

The results are derived under the assumption that the marginal income from general practice is higher than the wage rate in community health service:

$$(1) \quad w < \frac{q}{t}$$

We form the Lagrangian:

$$L = qn + w\tau + v(T - nt - \tau)$$
$$-\mu(n - n^{D})$$
$$-\gamma(\Phi - \tau)$$

Necessary and sufficient conditions for n > 0 and $\tau \ge \Phi$ to solve the problem are that there exists non-negative μ og γ such that:

(2)
$$\frac{\partial L}{\partial n} = 0$$

(3)
$$\frac{\partial L}{\partial \tau} \le 0$$
, $\tau = \Phi$ or $\frac{\partial L}{\partial \tau} = 0$

We have:

(4)
$$\frac{\partial L}{\partial n} = q - v'(T - nt - \tau)t - \mu = 0$$

(5)
$$\frac{\partial L}{\partial \tau} = w - v'(T - nt - \tau) + \gamma \le 0$$

We distinguish between constrained and unconstrained GPs.

The unconstrained GP

This GP does not experience patient constraints, implying $\mu = 0$. With $\mu = 0$, (4) can be written:

(6)
$$v'(T - nt - \tau) = \frac{q}{t}$$

The rhs of (6) expresses the income per time unit of adding a new person to the list. The unconstrained GP adjusts the number of patients such that the marginal income equals the

marginal utility of leisure.

We may now derive the unconstrained GP's labour supply to community health service. We show that assuming a non-binding minimum requirement constraint contradicts (1). Inserting $\gamma = 0$ and substituting $\frac{q}{t}$ for $v'(T - nt - \tau)$ in (5) we get:

(7)
$$w - \frac{q}{t} = 0$$
, which contradicts (1)

We thus conclude:

$$\gamma > 0$$
 and $\tau = \Phi$

Hence, the unconstrained GP does not contract for more than the required number of hours with the community health service.

The constrained GP

This GP experiences patient constraints, implying $\mu > 0$. With $\mu > 0$, (4) can be written:

(8)
$$v'(T - nt - \tau) + \frac{\mu}{t} = \frac{q}{t}$$

From (8) we see that a constrained GP has a marginal utility of leisure which is lower than the marginal income in general practice, $v'(T-nt-\tau) < \frac{q}{t}$. This is a constrained optimum, implying that the physician would be willing to fill a part-time position in community health service that would otherwise seem unprofitable.

Now we distinguish between two types of constrained GPs. Constrained type 1 has a binding minimum requirement constraint while constrained type 2 is not constrained by the municipality's minimum requirement. We show that a binding minimum requirement constraint, $\gamma > 0$, occurs when the marginal utility of leisure at $l = T - n^D t - \Phi$ is sufficiently high, $v'(T - n^D t - \Phi) > w$, whereas a sufficiently low marginal utility of leisure, $v'(T - n^D t - \Phi) < w$, results in a non-binding minimum requirement constraint, $\gamma = 0$.

Constrained type 1

A binding minimum requirement constraint implies $\gamma > 0$, $\tau = \Phi$ and $\frac{\partial L}{\partial \tau} < 0$

From (5) we have:

$$(9) w - v'(T - n^{D}t - \Phi) < -\gamma$$

Implying (10)
$$v'(T-n^Dt-\Phi) > w$$

One may refer to the marginal utility of leisure at $l = T - n^D t - \Phi$ as the reservation wage for voluntary involvement in community health service, and we define:

$$(11) w \equiv v'(T - n^D t - \Phi)$$

Differentiating (11) we get:

$$\frac{d\underline{w}}{dn^d} = -v''(l)t > 0$$

We see that this reservation wage is positively related to the GP's list size, which means that, ceteris paribus, a GP with a long list has a higher reservation wage than a GP with a short list.

Constrained type 2

A non-binding minimum requirement constraint implies $\gamma = 0$, $\tau \ge \Phi$ and $\frac{\partial L}{\partial \tau} = 0$

From (5) we have:

$$(12) \quad v'(T - n^D t - \tau) = w$$

Implying

$$(13) \tau > \Phi \Leftrightarrow v'(T - n^D t - \Phi) < w$$

We see that constrained type 2 GPs adjust the amount of hours they contract with the community health service such that the marginal utility of leisure equals the wage rate in community health service. Differentiating (12) we get:

$$\frac{d\tau}{dn^D} = -t < 0$$

In optimum, a marginal increase in list size results in a decrease in the hours contracted with the community health service.

REFERENCES:

Amemiya T. Tobit Models: A Survey. J Econometrics 1984; 24(1-2): 3-61.

Baker LC, Phillips KA, Haas JS, Liang SY and Sonneborn D. The Effect of Area HMO Market Share on Cancer Screening. *Health Services Research* 2004; 39(6): 1751-1772.

Baltagi BH, Bratberg E and Holmås TH. A panel data study of physicians' labour supply: the case of Norway. *Health Economics* 2005; **14** : 1035–1045.

Barros PP, Olivella P. Waiting lists and patient selection. *Journal of Economics & Management Strategy* 2005; **14(3)** : 623–646.

Bir A, Eggleston K. Physician Dual Practice: Access Enhancement or Demand Inducement? http://ase.tufts.edu/econ/papers/200311.pdf [2003].

Cragg JG. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica* 1971; **39**: 829-844.

Culler SD, Bazzoli GJ. The Moonlighting Decisions of Resident Physicians. *Journal of Health Economics* 1985; **4**: 283–292.

Eggleston K, Bir A. Physician dual practice. *Health Policy* 2006; **78**: 157–166.

Elstad JI. Flere leger, store bruk? Artikler om bruk av allmennlegetjenester, INAS rapport 1991:11. Institutt for sosialforskning, Oslo. [In Norwegian]

Godager G, Lurås H. I skyggen av fastlegeordningen: Hvordan har det gått med det offentlige legearbeidet? Hero rapport 2005:6, Helseøkonomisk forskningsprogram ved Universitetet i Oslo (HERO). [In Norwegian]

González P. Should physicians' dual practice be limited? An incentive Approach. *Health Economics* 2004; **13** : 505–524.

Hsiao C. Analysis of Panel Data. 2nd ed. Cambridge University Press: New York, 2003.

Iversen T. The effect of a private sector on the waiting time in a national health service. *Journal of Health Economics* 1997; **16** : 381-396.

Iversen T. The Effect of Patient shortage on General Practitioners' Future Income and List of Patients. *Journal of Health Economics* 2004; **23**: 673–694.

Iversen T, Lurås H. Economic Motives and Professional Norms: The Case of General Medical Practice. *Journal of Economic Behavior and Organization* 2000; **43**: 447-470.

Iversen T, Lurås H. The Importance of Micro-data for Revealing Income Motivated Behaviour Among GPs. In *Individual Decisions for Health*, Lindgren B (editor) London: Routledge, 2002.

Jan S, Bian Y, Jumpa M, Meng Q, Nyazema N, Prakongsai P, Mills A. Dual job holding by public sector health professionals in highly resource-constrained settings: problem or solution? *Bulletin of the World Health Organization* 2005; **83(10)**: 771-775.

Kristiansen IS, Mooney G. The general practitioners use of time: is it influenced by the remuneration system? *Social Science and Medicine* 1993; **3**: 393-399.

Langwell KM. Factors affecting the incomes of men and women physicians: further explorations. *Journal of Human Resources* 1982; **17**: 261-276.

Maddala GS. *Limited dependent and qualitative variables in econometrics*. Cambridge University Press: Cambridge (GB); 1983.

Melenberg B, Van Soest A. Parametric and Semi-Parametric Modelling of Vacation Expenditures. *Journal of Applied Econometrics* 1996; **11(1)**: 59-76.

Moffit RA. New developments in econometric methods for labour market analysis. In *Handbook of Labor Economics*, Ashenfelter OC, Card D (editors). Amsterdam: Elsevier, 1999; (3) 1367-1397.

Mroz T. The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. *Econometrica* 1987; **55** : 765-800.

Paulsen B, Fjermestad T. Populasjonsansvar i allmennpraksis. Tre års erfaringer med fastlegeordning i Norge. *Tidsskrift for den Norske Lægeforening* 1996; **116**: 2573-2576, (Summary in English) [In Norwegian]

Rizzo JA, Blumenthal D. Physician labor supply: Do income effects matter? *Journal of Health Economics* 1994; **13**: 433-453.

Shishko R, Rostker B. The Economics of Multiple Job Holding. *American Economic Review* 1976; **66(3)**: 298-308.

Showalter MH, Thurston NK. Taxes and labor supply of high-income physicians. *Journal of Public Economics* 1997; **66**: 73-97.

Statistics Norway. Standards for categorizing Norwegian municipalities. 1994.

Sæther EM. Physicians' Labour Supply: The Wage Impact on Hours and Practice Combinations *LABOUR* 2005; **19(4)**: 673–703.

Tobin J. Estimation of Relationships for Limited Dependent Variables. *Econometrica* 1958; **26:** 24-36.

Yip WC. Physician response to Medicare fee reductions: changes in the volume of coronary artery bypass graft (CABG) surgeries in the Medicare and private sectors *Journal of Health Economics* 1998; **17** : 675–699.

SERVICE MOTIVES AND PROFIT INCENTIVES AMONG PHYSICIANS¹

Geir Godager

Institute of Health Management and Health Economics
University of Oslo
P.O. Box 1089 Blindern,
NO-0317 Oslo, Norway
<g.c.h.godager@medisin.uio.no>

Tor Iversen

Institute of Health Management and Health Economics
University of Oslo
P.O. Box 1089 Blindern,
NO-0317 Oslo, Norway
<tor.iversen@medisin.uio.no>

Ching-to Albert Ma

Department of Economics
Boston University
Boston, Massachusetts, USA, and
Department of Economics
University of Oslo, Norway
<ma@bu.edu>

103

¹ This article is published in *International Journal of Health Care Finance and Economic*, 2009, s **9:1**,p. 39-57. There should only be minor typograhical differences between versions.

Abstract

We model physicians as health care professionals who care about their services and monetary rewards. These preferences are heterogeneous. Different physicians trade off the monetary and service motives differently, and therefore respond differently to incentive schemes. Our model is set up for the Norwegian health care system. First, each private practice physician has a patient list, which may have more or less patients than he desires. The physician is paid a fee-for-service reimbursement and a capitation per listed patient. Second, a municipality may obligate the physician to perform 7.5 hours per week of community services. Our data are on an unbalanced panel of 435 physicians, with 412 physicians for the year 2002, and 400 for 2004. A physician's amount of gross wealth and gross debt in previous periods are used as proxy for preferences for community service. First, for the current period, accumulated wealth and debt are predetermined. Second, wealth and debt capture lifestyle preferences because they correlate with the planned future income and spending. The main results show that both gross debt and gross wealth have negative effects on physicians' supply of community health services. Gross debt and wealth have no effect on fee-for-service income per listed person in the physician's practice, and positive effects on the total income from fee-for-service. The higher income from fee-for-service is due to a longer patient list. Patient shortage has no significant effect on physicians' supply of community services, a positive effect on the fee-for-service income per listed person, and a negative effect on the total income from fee for service. These results support physician preference heterogeneity.

1. Introduction

Economic theory is largely based on a hypothesis of self-interest. To a large extent, it is argued, many social phenomena may be explained as outcomes of interactions between selfish economic agents. Nevertheless, the selfish economic agent hypothesis is a simplifying assumption. Economists do recognize that even their own behaviors are not entirely consistent with self-interest, and that many social phenomena cannot be easily explained by it.

The self-interest hypothesis is probably unpalatable when it is applied to the health care market. There are serious frictions in the health care market due to hidden information and hidden action. One wonders why the complete collapse of the health market had not already occurred if physicians and health care professionals were completely guided by their selfish goals. In fact, Arrow (1963), in his seminal discussion of the medical market, already has called for a broader perspective. He also points out as a matter of fact that health care professionals are strongly influenced by ethical conduct, standards of care and service motives.

The literature on physician response to incentives is very large, and a major area in the entire health economics field.² The focus of this literature is often on the magnitude of empirical outcomes of an incentive innovation. For example, if the US Medicare system changes its fee structure, will physicians who experience an income loss perform more coronary artery bypass grafting procedures (Yip 1998)? Or, do obstetricians perform more cesarean sections when they are in financially less rewarding markets (Gruber and Owings 1996)? It is of course important to

-

² We cite only two papers here, but health economists will agree with us about this assertion.

study the direct effect of financial incentives on medical treatment. Physicians, however, perform services that may not generate the most monetary rewards for them. Arrow's "broader" perspective calls for studying effects of incentives on the many tasks that physicians perform.

In this paper we model physicians as health care professionals who care about their community services. Their preferences are a combination of community service and monetary rewards. Furthermore, we let these preferences be heterogeneous; different physicians trade off the monetary and service motives differently. Heterogeneity is an important assumption because preferences on monetary and service motives determine how physicians react to incentive schemes. Those physicians who care more about monetary rewards react more strongly to financial incentives than those who do not.

We set up a theoretical model for physician services in Norway. Various components of the model reflect the Norwegian health care system. There are two important elements in the description of the private practice physicians in Norway. First, each private practice physician has a list of patients under his care, and this list may have more or less than the number of patients he desires. The physician is paid a fee-for-service reimbursement together with a capitation per patient in his practice list.

Second, each physician is obligated to perform some community service in the municipality where he works. In fact, a municipality has the power to request 7.5 hours per week of community service from a physician. Physicians are paid an hourly wage for their community services. This hourly wage is quite low compared to the equivalent earning a physician can make

in private practice. This is the basis for our assumption that physicians are motivated by their preferences for community services to the municipalities. Despite a smaller financial reward, some physicians actually work more than the legally required amount of community service.

We assemble a data set on an unbalanced panel of 435 physicians, with 412 physicians for the year 2002, and 400 for 2004. The information includes physician personal characteristics, their community involvements, and private practices. Our estimations identify the effect of physician characteristics on their private practice styles as well as their community services. We look at services provided by physicians to their patients. Are they affected by whether the physicians think that they have enough patients in their lists? Does patient shortage affect physicians' supply of community health service?

We use a physician's amount of wealth and debt in previous periods as proxy for the physician's preferences for community service. First, for the current period, accumulated wealth and debt are predetermined. Second, wealth and debt likely capture lifestyle preferences because they correlate with the planned future income and spending. The actual implementation will use gross wealth and gross debt in the regressions. The higher is gross debt, the higher the future income required to pay for the interest. This likely means that the physician is less interested in providing community service, which is financially less rewarding.

In our study, physicians' community health service supply decisions are censored because municipalities may impose upon physicians up to 7.5 hours of work per week. When the dependent variable is censored, a linear regression model will give inconsistent estimates.

Instead, we estimate a random-effects tobit model on physicians' community service supply. For estimating the effects of indicators of service motive on the physicians' private practice service supply, we use a standard random-effects model, which controls for unobserved heterogeneity in our panel data.

The main results show that both gross debt and gross wealth have negative effects on physicians' supply of community health service. Gross debt and wealth have no effect on feefor-service income per listed person in the physician's practice, and positive effects on the total income from fee-for-service. The higher income from fee-for-service is due to a longer patient list. Patient shortage has no significant effect on physicians' supply of community services, a positive effect on the fee-for-service income per listed person, and a negative effect on the total income from fee for service.

Our results suggest that policies may affect physicians differentially. A mandatory increase in the hours of community service municipalities may impose on physicians likely will be binding on the majority of physicians, but may not be so for those physicians with lower gross wealth and debt. On the other hand, an increase in the remuneration of community services may relax this constraint for the majority of physicians, but have little effect on those with lower gross wealth and debt.

The paper proceeds as follows. Section 2 describes the study setting and reviews the literature. We set up a model in section 3, and derive a set of hypotheses. Section 4 presents the

data and descriptive statistics. In Section 5 the strategy for empirical analysis is explained and results are presented. Concluding remarks are given in Section 6.

2. Study setting and literature review

We use data from Norway in this study. Norway is a country of about 4.5 million inhabitants. Norwegians' health care is covered by a national health service, which is mainly tax-financed. Hospitals are publicly owned, and inpatient care is free to users. Outpatient consultations with primary care physicians and specialists are offered respectively with a copayment of about US\$25 and US\$40 in 2006. Since the implementation of the Regular General Practitioner Scheme in 2001, each inhabitant of Norway has been listed with a General Practitioner (GP), or primary care physician. About 90% of GPs are self-employed, private physicians contracting with municipalities, with the remaining GPs employed by the municipalities. Each GP has a list of patients. In 2004 the average list-size was between 1250 and 1300 people. Besides providing primary care, GPs act as gatekeepers. A referral by a GP is required for consultations with health care specialists. The national insurance covers all expenditures if copayments for physician services and medicines within a year exceed a deductible of about US\$250.

The Regular General Practitioner Scheme of 2001 required each inhabitant to submit to the National Insurance Administration up to three preferred physicians. GPs submitted to the Administration the maximum number of patients they were willing to include in the practice list. A matching process respecting patient and GP preferences formed the GP patient lists. For many physicians the maximum number of patients they were willing to accept exceeded the number of

people who showed interest in being listed with them. The administration then allocated inhabitants who did not submit any physician preference (30 percent of the adult population) to these GPs. As of June 2001, after this second round of assignments, about 30 percent of the GPs still had at least 100 patients less than the number of patients they were willing to take. In the paper we say that these GPs experience a shortage or deficit of patients.

Private practice general practitioners have three sources of revenue. First, there is a fee-for-service payment; a GP provides various services to patients in return for a fee from the national insurance. Second, for each consultation, a GP receives a copayment from the patient. Third, a GP receives a capitation fee from the municipality in which he serves. The capitation amount is based on the number of listed patients with the GP without any risk adjustment. Each of the three components constitutes about one third of the income of an average practice.

In Norway preventive health care at childcare centers and schools, and regularly medical care at nursing homes and prisons are served by GPs working part-time in municipalities. Such community health services are remunerated according to a fixed salary scheme that is negotiated between the state and the Norwegian Medical Association. The community service remunerations are in terms of hourly wages and tend to be lower than the equivalent rates in private practice. GPs are also entitled to a "practice compensation" to cover costs in their practice while working for the municipality, and it is paid on an hourly basis. In Godager and Lurås (2005) the remuneration rate for community service is estimated to be between 38% and 66% of the equivalent private practice rate. This range is due to variations in cost reductions in GPs' private practice while working for the municipality. According to current regulations, a

municipality can require GPs to perform up to 7.5 hours of community services per week. A municipality is obliged to strive for an equitable distribution of community health workload among the GPs if they choose to enforce the regulation. In nearly all municipalities at least one GP works more that 7.5 hours of community work, so it seems that those who work less than 7.5 hours would not have preferred to work more, but may well have preferred to work less.

Community service is provided in normal work hours and does not substitute for leisure. A physician may have to be absent from his private practice one day a week to provide services at nursing homes and childcare centers. We have found a negative correlation between the private practice list size and community service hours; this is consistent with community services using up physicians' time that would otherwise be available for private practice. Generally, physicians should find it more rewarding to build reputation through superior services in private practices. We regard community services as activities mainly motivated by nonprofit-seeking objectives.

Several papers have studied the impact of economic incentives since the health system reform in Norway. Iversen (2005) studies whether patient shortage will lead a GP to increase services provided to patients in the practice. The study shows that GPs with patient shortage in fact compensate for their lower capitation payment by earning more fee-for-service incomes. Carlsen and Norheim (2003) investigate whether the patient list system has influenced general practitioners' self-perception as gatekeepers. They find that GPs generally have become less concerned with the gatekeeper role. Rather, GPs believe that they should provide better services to keep patients from switching to other physicians.

In Lurås (2007) a nationally representative sample of Norwegians are surveyed about satisfaction with their GPs. She finds that if a patient's GP has a patient shortage, then she is likely to be dissatisfied in most quality dimensions except waiting time. Iversen and Lurås (2008) add to this result by supplementing the earlier study by registrar data. They find that patients of GPs with patient shortage tend to switch GPs more often, even though these GPs already provide more services.

Using cross sectional data from 2002, Godager and Lurås (2006) study the effect of patient shortage on GPs' supply of community health service. From tobit regressions, they find that GPs experiencing a patient shortage contract for more hours of community health service. The dataset used in Godager and Lurås (2006) is the same as the 2002 part of the data in this paper.

We are unaware of any paper that studies the relationship between physician indebtedness and physicians' service decisions. There are, however, some papers that study the effect of study loans on physicians' occupational choices. Fox (2003) finds that physicians who have had large study loans are less likely to enter academic medicine, which is financially less rewarding.

Bazzoli (1985) and Thornton (2000) find that medical students' magnitude and types of loans have an impact on physician specialty choices. Culler and Bazzoli (1985) study factors that affect resident physicians taking a second job; when making moonlighting decisions, residents are influenced by debt and other economic factors.

3. The model

We present a model of physician decision on private practice and community services. A physician has a private practice, where he provides services for patients who are enrolled with him. The physician also spends some time to work for the municipality. We call this community service. While GPs' work in private practice usually belongs to the discipline of general medicine, community services at the municipality typically are on nursing home care, prisons, vaccination for school children, administrative work, and related community medicines. The contract between the physician and the municipality stipulates that a minimum number of hours of community service may be required.

The physician receives two kinds of payments for treating patients at his private practice. First is the patient list component of the revenue. The physician receives a capitation payment, a lump sum per patient who has elected to be in the doctor's practice. Second is the fee-for-service component of the revenue. The physician receives a payment based on the service that is provided to a patient.³ Community services are also remunerated, and they are paid on an hourly basis.

The payment for a unit of private practice service, s, is denoted by α ; the community service has an hourly remuneration rate β . While the fee-for-service rate α is based on the quantity of services, we will interpret α as an equivalent hourly rate, so that the private-practice and community-service remuneration rates are comparable. Alternatively, we may interpret s as hours of private practice. The remuneration rate for community service is lower than private

_

³ Physicians also receive copayments from patients for office consultations, but we will ignore this revenue source.

service, so we assume that $\alpha > \beta$. The last component of payment is the capitation rate per patient enrolled in a physician practice; this is denoted by γ .

Let *n* denote the number of patients who are enrolled in the physician practice, and *s* the service that the physician supplies to a patient. Let *a* denote the amount of community service the physician provides at a municipality. The physician decides on these three variables subject to various constraints to be explained below.

The physician incurs a total cost of C(ns + a) when he provides s units of services to each of n patients, and when he supplies a units of community service. The cost function includes both the physician's time cost and other necessary input costs for providing s services to each of n patients, and the community service a. For convenience, we have chosen to let cost be a function of the sum of private and community services. The function C is increasing and convex. We will also assume that it is twice differentiable, and that the marginal cost (first order derivative) increases without bound. The physician derives utility $\theta V(a)$ from community service a. The function V is an increasing and concave function, and θ a positive parameter. We postulate that the physician is motivated to provide community service, 4 and this motivation is captured by the utility $\theta V(a)$. We will discuss how we proxy for the preference parameter θ .

For simplicity, we have assumed that the physician's concern for patients in his private practices is purely motivated by profits. This may not seem entirely consistent with the assumption that physicians derive a utility from serving the community besides the monetary

-

⁴ We assume that voluntary and involuntary community services lead to the same utility. A more general form of the utility from community service is $V(\theta, a)$, but our results are unaffected by the simpler form $\theta V(a)$.

remuneration. In the appendix, we have examined the robustness of the model. There we allow the physician to derive a utility from serving patients in his private practice, and show that the predictions by the model remain valid.

There are two constraints that restrict the physician's choice of the number of patients in his practice, as well as the service for each patient. First, we let D be the maximum number of patients that the physician can have. This maximum demand D is assumed to be exogenously given.⁵ In a short period of time, the physician cannot influence the total number of patients willing to be listed with him. Nevertheless, the physician may decide to serve less than D patients. Therefore, the first constraint for the physician is $n \le D$.

In the absence of this constraint, a physician may want to enroll more patients. If indeed the physician does want a larger patient list, the constraint will become binding (n = D), in which case we say that the physician has a shortage of patient or that he is rationed. We will not impose a minimum community service constraint now. The basic model will be used later for studying this possibility.

The second constraint concerns the physician's service intensity. We assume that the service per patient, s, is limited to a range $[S_1, S_2]$, with $S_1 < S_2$. This range of services describes the physician's control on patients, or the extent of physician agency. Superior medical knowledge and experience allow the physician to dictate to some extent the services patients receive.

⁵ Using Norwegian data, Iversen and Lurås (2008) show that service intensity has a negative impact on the number of patients switching physicians, but the magnitude of the response is too small to be of importance.

Variations in services, however, are subject to some limits. We bound these variations by an interval. We assume that S_1 and S_2 are exogenous. Within this range, the physician is able to dictate the service to the patient: $S_1 \le s \le S_2$.

Given the payment parameters, fee-for-service rate α , community service rate β , and capitation rate γ , if the physician has n patients in his practice, and provides s services to each patient, as well as community service a, his payoff is

(1)
$$U(s,n,a) = \alpha sn + \beta a + \gamma n + \theta V(a) - C(ns+a).$$

The utility function in (1) contains the financial rewards from private practice and community service (the first three terms), an enjoyment from serving the municipalities, and the cost of services. The physician's behavior is described by his choice of n, s and a that maximize his utility in (1) subject to the constraints $n \le D$ and $S_1 \le s \le S_2$.

We begin by considering cases when the constraint $n \le D$ does not bind. Here, the physician is not rationed and can choose the optimal number of patients for his practice without worrying that insufficient patients will elect to join. The first-order condition of U with respect to n is

(2)
$$\frac{\partial U}{\partial n} = s \left[\alpha - C'(ns + a) + \frac{\gamma}{s} \right] = 0$$

when the constraint $n \le D$ does not bind. Now consider the first-order derivative of U with respect to service s:

(3)
$$\frac{\partial U}{\partial s} = n \left[\alpha - C'(ns + a) \right] < 0.$$

From the first order condition (2), the first-order derivative with respect to s in (3) must be negative. This implies that the optimal value of s is S_I , the lower bound on the range of service.

Having an extra patient entitles the physician to obtain the capitation payment. The physician cares about total service *ns*. By reducing *s* and raising *n* to keep *ns* constant, the physician already raises his payoff due to the capitation payment. When there is no patient shortage, the physician tends to provide less service and enrolls more patients.

We have not included a utility component in the physician's service in the private practice. Such a utility may tend to raise the value of s in the above calculation. The tendency to increase n due to capitation remains robust for many utility specifications; see the Appendix.

Next, we differentiate the objective function U with respect to community service a:

(4)
$$\frac{\partial U}{\partial a} = \beta + \theta V'(a) - C'(ns + a).$$

From (3), and the assumption that $\alpha > \beta$, the expression in (4) must be strictly negative when θ is sufficiently small. Community service has a lower remuneration ($\alpha > \beta$). If the physician does not value community service sufficiently, he chooses the minimal level.

Now we consider the case when the constraint $n \le D$ binds. Here the first-order derivative of U with respect to n is positive at n = D:

$$\frac{\partial U}{\partial n} = s \left[\alpha - C'(Ds + a) + \frac{\gamma}{s} \right] > 0.$$

The first-order derivative with respect to s is

$$\frac{\partial U}{\partial s} = D[\alpha - C'(Ds + a)].$$

If D is small, then the first-order derivative evaluated at $s = S_1$ will likely be positive and the optimal s is strictly bigger than S_I . In fact, the first-order derivative may remain positive for all service levels, so that we may have a corner solution $s = S_2$. In such an equilibrium the community service a will be decreasing in D. For an interior solution, s is in $[S_1, S_2]$, and will be given by setting the above first-order derivative to zero. Finally, the first-order derivative (4) applies, and for an equilibrium where a > 0, it will be set at zero.

When the constraint $n \le D$ binds, and when the physician picks a service per patient in the interior of $[S_1, S_2]$, we can use the first order conditions:

$$\left[\alpha - C'(Ds + a)\right] = 0$$

$$\beta + \theta V'(a) - C'(Ds + a) = 0$$

to obtain comparative static results. At the service intensity interior solution, the equilibrium community service a is increasing in the preference parameter θ , but does not vary with the rationed list size D while the equilibrium service s is decreasing in D. 6

A physician having stronger preferences for community services will cut back more on private practice. This is because community services raise the marginal cost of supplying services to patients. Finally, a higher value of θ implies a larger supply of community service.

The community service parameter θ may capture physicians' preferences on lifestyle and work over the long term. Our model can be regarded as a component of a physician's dynamic decisions on his private practice and community services. The decision variables are current

⁶ Use the two first-order conditions to eliminate the term C' to get $\alpha = \beta + \theta V'(a)$. Hence, given an interior solution of s a change of community service a is only related to θ .

choices while the parameters capture earlier decisions such as wealth and debt accumulations. For the empirical implementation, we proxy the community service parameter θ by gross wealth and gross debt under the assumption that θ is decreasing in these variables.

A physician who does not value community services highly may prefer luxurious consumption, which often takes the form of durables. Such a physician likely accumulates more wealth and debt in his financial portfolio. Perhaps more important, his higher debt obligation cannot afford him the "luxury" of performing community service, which has a much lower remuneration than in private practice.⁷ These observations are consistent with our assumption that θ is decreasing in wealth and debt.

On the other hand, luxurious consumption may be on nondurable. A physician with this aptitude may not accumulate high wealth and debt, although he prefers to perform little community service. The behavior of such a physician is less agreeable with our assumption, which would identify him as one with a high value of θ . Our empirical results then would not yield any significant effects. Hence, our assumption that θ is decreasing in wealth and debt is conservative.

To summarize, we list several predictions of our model:

1. Physicians who have patient shortage tend to supply more service per patient; conversely, physicians who have no patient shortage tend to supply less service per patient.

-

⁷ Furthermore, the amount of wealth and debt in a financial portfolio likely depends on a person's lifecycle, which is controlled by the age variables.

- 2. Physicians' community service does not depend on the list size when they face a patient shortage and when the optimal service per patient is an interior solution.
- 3. With both patient shortage and constrained service per patient, the physicians' community service is decreasing in the rationed list size.
- 4. The stronger are physicians' preferences for community service, the larger the amount of community service they supply and the shorter the preferred list of patients. When physicians' gross wealth and debt are negatively related to their preferences for community services, physicians who have accumulated higher gross wealth and debt perform less community services.

4. Data and descriptives

A survey of 35 Norwegian municipalities and two districts of the city of Oslo form the basis of the data for analysis. This survey was initiated by us and put together by municipality administrative staff. The data contain information of physicians who participated in community health services at the said municipalities and districts for the years 2002 and 2004.

The municipalities and Oslo districts in the survey were randomly selected within groups stratified according to geography and a measure of centrality according to the classification by Statistics Norway (Norwegian Official Statistics, 1999). The stratification aims to obtain a representative sample of Norwegian municipalities. In 2002, all municipalities responded to the survey, while in 2004, four municipalities failed to respond (a corresponding 89% response rate).

120

,

The classification assigns each municipality to one of four groups based on travel time from the municipality to the nearest densely populated area.

The four municipalities that did not respond were small, and so were the numbers of physicians in these municipalities relative to the total.

The survey data were merged with registrar data from the Norwegian primary physician database, which describes characteristics of each GP and each GP's patient list. GP characteristics include age, gender, number of children according to age groups, taxable income, wealth and debt. The GP practice characteristics include preferred numbers of patients, actual number of patients according to gender and age, and the total fees from national insurance.

Primary care physicians who did not provide any community service were not in the survey. The municipalities simply did not register these physicians in their administrative files. Those physicians in the registrar data who did not appear in the survey were assigned zero hours of community service in the corresponding municipalities or Oslo districts.

For confidentiality and privacy protection, each physician in the survey was informed and given the opportunity to withdraw participation from the survey. No such request was received and the merged data from the 2002 survey was made available for research four months after data collection. The merged data from the 2004 survey was available for the researchers eight months after data collection.

The data set is an unbalanced panel of 484 physicians. There were 466 physicians for the year 2002, and 440 in the year 2004. We exclude GPs who contract with more than one

9 The Norwegian primary physician database is administered by the Norwegian Social Science Data Service (NSD)

and provides information of individual GPs.

121

municipality (6 physicians each year) because we are unable to disaggregate their total practice income into the municipality sources. We also exclude salaried GPs (28 physicians in 2002 and 22 physicians in 2004) because their economic incentives are different from the private GPs who contract with a municipality. Then we exclude those GPs who were both salaried and contracted with more than one municipality (1 physician each year not in the previous exclusions). In the primary physician registrar, information of *Annual income from fees from national insurance* or *Gross debt* and *Gross wealth* was missing for 19 physicians in 2002 and for 11 physicians in 2004. Our analysis is then based on data of a total of legitimate 812 observations (412 in 2002 and 400 in 2004) of 435 GPs.

Table 1 presents descriptive statistics of the full panel. The last two columns decompose the total variation into 'between physician' (b) and 'within physician' (w) variation. On average a physician works 4.88 hours per week of community health services, with a maximum of 22.5 hours per week. The between variation as a proportion of total variation is 71 percent and accordingly, the within variation is 29 percent of the total variation. About 14 percent of the GPs work more than the 7.5 hours per week, which is the legal requirement that a municipality may impose on GPs. On average a GP's preferred list size (1393) is slightly larger than the actual list size (1316). While 22 percent of the GPs experience a shortage of patients, 8 percent have a list larger than they prefer.

_

 $^{^{10}}$ While 'between physician variation' measures the variation in physician averages, 'within physician variation' measures the variation around the average of the two periods for each physician.

Table 1 Descriptive statistics for the panel

Variable	Definition	Mean	Std. Dev.	Min	Max	b	W
Total-hour	Total hours per week in						
	community health service	4.88	4.27	0	22.5	0.71	0.29
Volunt-hour	Binary variable set to 1 if Total-						
	hour>7.5, otherwise 0	0.14		0	1	0.65	0.35
Prefer-list	The GP's preferred list size	1393	378	100	2500	0.76	0.24
List	Actual list size	1316	383	98	2798	0.79	0.21
Prop-female	Proportion of females on list	0.51	0.10	0.25	0.86	0.92	0.08
Prop-old	Proportion of 70 and older on						
•	list	0.11	0.06	0.00	0.37	0.88	0.12
Shortage	Binary variable set to 1 if						
	(Prefer-list – list)>100,						
	otherwise 0	0.22		0	1	0.54	0.46
Many	Binary variable set to 1 if						
	(Prefer-list – list)<-100,	0.00			1	0.50	0.50
T . I PPG	otherwise 0	0.08		0	1	0.50	0.50
Total-FFS	Annual income (NOK) from fees	550100	205717	875	2702649	0.70	0.00
FFS-NI	from national insurance Annual income (NOK) from fees	558102	285717	013	2702049	0.70	0.30
FFS-NI	from National insurance per						
	listed person	440.38	251.24	0.84	3677.85	0.47	0.53
Gr-debt	Gross debt in million NOK	1.15	1.08	0.04	6.86	1.00	0.00
Gr-wealth	Gross wealth in million NOK	1.13	0.92	0	8.21	1.00	+
	Net wealth in million NOK			-5.32			0.00
Net-wealth		0.06	1.36	-5.52	4.79	1.00	0.00
Gen-Med	Binary variable set to 1 if GP						
	specialist in general medicine, otherwise 0	0.59		0	1	1.00	0.00
Comm-Med	Binary variable set to 1 if GP	0.57		U	1	1.00	0.00
Comm-mea	specialist in community						
	medicine, otherwise 0	0.06		0	1	1.00	0.00
Mid-age	Binary variable set to 1 if					1100	0.00
mu uge	40 <gp's 55<="" age="" td="" ≤=""><td>0.57</td><td></td><td>0</td><td>1</td><td>1.00</td><td>0.00</td></gp's>	0.57		0	1	1.00	0.00
Old-age	Binary variable set to 1 if GP's						
O	age > 55	0.17		0	1	1.00	0.00
Male	Binary variable set to 1 if GP is						
	a male, otherwise 0	0.74		0	1	1.00	0.00
Married	Binary variable set to 1 if GP is						
	a married, otherwise 0	0.78		0	1	0.89	0.11
Low-Central	Binary variable set to 1 if						
	municipality has lowest level of	0.04			1	1.00	0.00
11.1.0	centrality; otherwise 0	0.04		0	1	1.00	0.00
Med-Central	Binary variable set to 1 if						
	municipality has second lowest	0.07		0	1	1.00	0.00
II: al. 1	level of centrality; otherwise 0	0.07		U	1	1.00	0.00
High-1-	Binary variable set to 1 if municipality has second highest						
Central	level of centrality; otherwise 0	0.19		0	1	1.00	0.00
High-	Binary variable set to 1 if	0.17		0	1	1.00	0.00
піgn- Central	municipality has highest level of						
cemiai	centrality; otherwise 0	0.70		0	1	1.00	0.00

As described previously, a GP's total practice income consists of capitation fees (NOK 299 per person¹¹ listed in 2003), patient copayments and service fees from the national insurance. We do not have reliable data on patient copayments. However, as patient copayments and service fees from the national insurance are both proportional to the volume of services provided, we use the annual income from national insurance fees as a proxy for the fee-for-service income. From Table 1 the mean of this fee is NOK 558102 per physician per year.

Table 1 also displays the average physician debt and wealth. Gross wealth (*Gr-wealth*) is defined as the sum of real capital (including housing value) and financial assets (bank deposits and other financial assets). Gross debt (*Gr-debt*) is personal debt including mortgage balance. Net wealth (*Net-wealth*) is the difference between gross wealth and gross debt. The mean gross debt is 1.15 million NOK, while the mean gross wealth is 1.21 million NOK. Together these figures imply a positive average net wealth. The variation in the debt and wealth figures is considerable. Because we only have data on wealth and debt for the year 2002, the within physician variation is zero for these variables. The majority of GPs are between 40 and 55 years old, and 74 per cent of them are men. Seventy-eight per cent of GPs are married, and they have on average 0.27 children below six years of age. About 6 percent of physicians are specialists in community medicine, while 59 percent have earned a specialist degree in general medicine. From Table 1, 4 percent of the GPs practice in a municipality with the lowest level of centrality, while 70 percent practice in a municipality with the highest level of centrality.

-

¹¹ 1 USD was approximately 6.30 NOK in 2003.

Wealth and debt are measured at the individual level. A GP's decisions are likely influenced both by the spouse's wealth and debt, too. We would prefer to have access to household wealth and debt, but because this information is unavailable to us, we could only use a physician's marital status (*Married*) as a control. In auxiliary regressions we have introduced interaction terms between *Married* and *Wealth/Debt* to check whether marital status affects the impact of *Wealth/Debt*. These interaction terms have not yielded statistically significant effect, and we have dropped them (so they do not appear in Tables 3 and 4 below).

Table 2 contains the descriptive statistics according to physicians' involvement in community health services. We categorize the information according to whether the physicians work more or less than 7.5 hours, the obligation that municipalities may impose upon them.

Those physicians who work more than 7.5 hours may have chosen to do so voluntarily. Those physicians who work voluntary hours have shorter preferred lists and actual lists. However, the two groups of physicians share similar characteristics with respect to gender and elderly proportion in their patient lists. The proportion of GPs with patient shortage is higher among those who work voluntary hours of community health service (28 percent) than those who do not (21 percent)¹². GPs who work less than 7.5 hours have both higher gross debt and gross wealth, but those who work more than 7.5 hours have a higher net wealth. Finally, those who work voluntary hours at municipalities are more likely to be specialists in community medicine.

¹² Likewise, the proportion doing voluntary community health service conditional on *Shortage* is 17 percent, and the proportion doing voluntary community health service conditional on *Many* is 4 percent.

Gross wealth and gross debt are independent variables in the regressions. 13 As we have argued in the previous section, those physicians who have higher levels of gross wealth are likely to have a more affluent lifestyle, and those who have higher debt require more income to pay for finance charges and interests. We associate weaker preferences for community services with higher physician gross wealth and debt.

Table 2 Descriptive statistics according to physician community health service

	<i>Volunt-hour</i> = 0 (No. obs. = 700)		<i>Volunt-hour</i> = 1 (No. obs. = 112)					
Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
<i>m</i> . 1.1	0.00	0.00	0	7.5	40.00	4.00	0	00.5
Total-hour	3.60	2.62	0	7.5	12.83	4.02	8	22.5
Prefer-list	1412	377	100	2500	1274	362	300	2000
List	1336	385	98	2798	1187	349	212	2045
Prop-female	0.51	0.10	0.28	0.86	0.50	0.10	0.25	0.76
Prop-old	0.10	0.06	0.00	0.35	0.12	0.07	0.00	0.37
Shortage	0.21		0	1	0.28		0	1
Many	0.09		0	1	0.03		0	1
Total-FFS	562691	293449	875	2702649	529421	230673	32892	1344763
FFS-NI	435.65	255.99	0.84	3677.85	469.99	217.95	26.33	1527.60
Gr-debt	1.19	1.10	0	6.86	0.89	0.85	0	3.35
Gr-wealth	1.24	0.96	0	8.21	1.02	0.56	0.03	2.67
Net-wealth	0.06	1.41	-5.32	4.79	0.13	1.06	-2.37	2.36
Gen-Med	0.58		0	1	0.63		0	1
Comm-Med	0.04		0	1	0.20		0	1
Mid-age	0.54		0	1	0.71		0	1
Old-age	0.18		0	1	0.12		0	1
Male	0.73		0	1	0.79		0	1
Married	0.76		0	1	0.79		0	1
Low-Central	0.03		0	1	0.11		0	1
Med-Central	0.07		0	1	0.10		0	1
High-1-	0.19		0	1	0.19		0	1
Central								
High-Central	0.71		0	1	0.61		0	1

 $^{^{13}}$ Since Net wealth = Gross wealth - Gross debt, we could have used any two of the three measures in the regressions. We also have access to data on interest payment. The coefficient of correlation between Gross debt and interest payment is 0.93, so interest payment does not add any information.

5. Empirical specification and results

We would like to know what determines GPs' community services and private practice. The predictions of our model are summarized at the end of Section 3. The exogenous variables in the model are used as regressors. Hence, *Shortage*, *Many*, *Gr-debt* and *Gr-wealth* are included in order to test predictions from our theory. In addition, we include *Prop-female* and *Prop-old* to control for variation in list compositions. We also control for several physician characteristics, such as socio-demographic factors (*Mid-age*, *Old-age*, *Male* and *Married*) and type of specialty (*Gen-Med* and *Comm-Med*) because physicians who have chosen a specialty in community medicine likely provide more hours of community services. Finally, we adjust for the level of centrality of the municipality a physician practices in. For instance, distance to the nearest hospital is correlated with a municipality's centrality and possibly has an impact on private physician practice and community health service.

In our study, GPs' labor supply decisions on community service are censored because municipalities may impose up to 7.5 hours of work per week on each physician. Furthermore, in our data, we observe cases in which GPs work less than 7.5 hours. We do not know if individual GPs are experiencing an enforced minimum requirement, so we must allow the censoring threshold to vary between GPs. In other words, each physician faces his own censoring threshold. When the dependent variable is censored, a linear regression model will give inconsistent estimates (Tobin, 1958). Many tobit models have been developed to take account of a censored dependent variable, and such models are frequently used in labor econometrics (Moffit, 1999).

Different municipalities may want different numbers of hours of community services from GPs. A municipality might only let a GP work a fraction of the time of community service that GP would have preferred. This kind of rationing by municipalities on GPs' community service seems improbable in our setting. Survey data from Norway show that only three per cent of GPs working less than eight hours of community work would have preferred more. Furthermore, municipalities are required to strive for an equitable distribution of hours of community work among GPs. In practice, equitable distribution means that a GP should not be asked to work more than 7.5 hours if someone else with less than 7.5 hours of work would prefer more. In our data only in six municipalities did all GPs there work less than 7.5 hours; altogether, there were a total of 28 GPs (or six percent of the total) in these six municipalities. We continue with the assumption that whenever a physician is observed to have worked less than 7.5 hours, it is a censored observation.

Let \tilde{y}_{ii} denote the number of hours of community service GP i prefers to work in time period t; we regard \tilde{y}_{ii} as a latent variable. Further let y_{ii} denote the actual number of hours of community service GP i has provided in time period t. When y_{ii} is less than 7.5, we do not know if this is a result of the physician's choice or the municipality's imposition, and can only infer that $\tilde{y}_{ii} \leq y_{ii}$. In this case we say that the physician's community service supply has been censored. Again, note that the censoring threshold on y_{ii} is allowed to vary across physicians and periods. For $y_{ii} > 7.5$ we assume that the community service provided is the GP's own choice. We assign the individual specific thresholds in period t, c_{ii} , according to the following

rule: $c_{it} = 0$ when $y_{it} > 7.5$, and $c_{it} = y_{it}$ when $y_{it} \in [0, 7.5]$. Letting I_{it} denote an indicator variable equal to 1 if y_{it} is censored, and 0 otherwise, we specify our censored regression model:

$$y_{it} = (1 - I_{it})(\beta' x_{it} + u_i + \varepsilon_{it}) + I_{it} c_{it},$$

where β is a vector of parameters, and x_{ii} a vector of explanatory variables. The variable u_i denotes random effects and is assumed to be i.i.d. $N(0, \sigma_u)$ while ε_{ii} 's are residuals and are assumed to be i.i.d. $N(0, \sigma_\varepsilon)$ and independent of u_i . The estimation is by maximum likelihood in STATA 10. The main results of the estimation are in Table 3.

Table 3 The estimated effect of physician characteristics on hours of community health service. Random-effects tobit model.

	Total-hour
Prop-female	-3.62 (11.44)
Prop-old	23.54 (16.28)
Shortage	-0.24 (1.36)
Many	-4.84 (2.74)
Gr-debt	-2.53 [*] (1.01)
<i>Gr-wealth</i>	-3.70 [*] (1.59)
Gen-Med	1.58 (2.02)
Comm-Med	10.36** (3.30)
Mid-age	3.09 (2.46)
Old-age	-4.32 (3.65)
Male	2.08 (2.81)
Married	-0.28 (1.72)
Med-Central	-0.12 (4.77)
High-1-Central	-6.51 (4.14)
High-Central	-5.94 (3.79)
Constant	-4.80 (8.27)
ρ	0.92
No. left-censored observations	700
No. of observations	812
No. GPs	435
No. observations per GP	Min: 1
	Avg: 1.9
	Max: 2

Estimates with '*' ('**') indicate that the parameter is significantly different from zero at the five (one) percent level for a two-tailed test.

From Table 3, both *Gr-debt* and *Gr-wealth* have negative and statistically significant effects on GPs' total number of hours of community health service. These are according to the prediction of our model. The magnitude of these effects is large. An increase of 10 percent from the mean of *Gr-debt* and *Gr-wealth* (which results in no change in net wealth) is expected to decrease community service by about 0.6 hours, or 12 per cent of the mean number of hours worked. Being a specialist in community medicine (*Comm-Med*) contributes positively to community service, while a higher degree of centrality has a negative effect. Patient shortage (*Shortage*) has a statistically insignificant effect on GPs' supply of community service.

We use a standard parameter to measure the latent, physician-specific heterogeneity in the supply of community health service. This parameter, ρ , is defined as the ratio of the variance of the physician-specific effect to the variance of the 'gross disturbance' $u_i + \varepsilon_{ii}$, i.e., $\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2}$. The parameter has the alternative interpretation of the coefficient of correlation between two 'gross disturbances' from the same physician in different years. The value of ρ =0.92 indicates that the unobserved heterogeneity is significant. Accounting for physician heterogeneity in community health service supply is important for the estimation.

As a robustness check we run a binary random-effects logit model in which the dependent variable is set to one if the physician works more than 7.5 hours of community services and to zero otherwise. We find that the effects of gross wealth and gross debt come up with similar signs and levels of statistical significance as in the random-effects tobit model.

We are also interested in estimating the impact of the indicators of service motive on the provision of services in the physicians' private practices. Again, since we have panel data, we are able to account for unobserved heterogeneity in the estimation. We fit a standard model¹⁴ of the form:

$$z_{it} = \gamma' \mathbf{x}_{it} + \alpha_i + v_{it}$$
 $(i = 1, ..., 435; t = 1, 2),$

where z_{it} is the dependent variable for GP i in time period t, and \mathbf{x}_{it} a vector of explanatory variables. We will use the national insurance total income from fees, both average (with respect to list size) and total, as the dependent variables. The variable α_i is a GP-specific random variable that captures unobserved GP heterogeneity; this effect is constant over time. Finally, v_{it} denotes the residuals. We assume that:

- (a) $E(v_{it})=0$,
- (b) Var $(v_{it}) = \sigma_v^2$
- (c) Cov $(v_{it}, v_{is}) = 0$
- (d) $E(\alpha_i) = 0$
- (e) $Var(\alpha_i) = \sigma_{\alpha}^2$
- (f) $Co v(\alpha_i, v_{it}) = 0$

If the random effects model is valid, we must have $Cov(\alpha_i, x_{i,t}) = 0$. We test this restriction by a standard Hausman-test. ¹⁵ From Table 4 we see that the Hausman statistic is not statistically significant at the conventional five-percent level, so we proceed with the random effects model.

-

¹⁴ See for instance Cameron and Trivedi (2005), Ch. 21.

¹⁵ If the restriction is rejected, the fixed effects model is selected. In the fixed-effects model α_i cancels; hence, the model is robust. When they are valid, the random effects estimators are more efficient than the fixed effects estimators. In addition, we are able to test the effect of time-invariant variables.

Table 4 The estimated effect of physician characteristics on the total and per listed-patient fee-for-service incomes. Random-effects model with robust standard errors.

	FFS-NI	Total-FFS
Prop-female	390.73 (270.99)	501735 ^{**} (160284)
Prop-old	120.34 (345.11)	552532 ^{**} (198814)
Shortage	109.95 ^{**} (30.39)	-39452 [*] (18804)
Many	-5.68 (22.03)	65291 [*] (29382)
Gr-debt	12.10 (11.80)	51372 ^{**} (15246)
Gr-wealth	35.40 (19.30)	53598 ^{**} (16141)
Gen-Med	50.55 ^{**} (18.28)	101367** (30546)
Comm-Med	21.27 (44.67)	-72471 (42089)
Mid-age	-5.10 (27.93)	-20497 (36924)
Old-age	-13.69 (52.19)	-98268 [*] (46103)
Male	107.17 [*] (52.55)	164412 ^{**} (40224)
Married	10.95 (26.10)	21567 (25449)
Med-Central	-86.85 (64.48)	21207 (57088)
High-1-Central	-68.89 (65.33)	101764 (54427)
High-Central	-131.19 [*] (60.99)	53653 (45773)
Constant	144.19 (174.83)	-114233 (113115)
ρ	0.36	0.75
No. of observations	812	812
No. GPs	435	435
No. observations per GP	Min: 1	Min: 1
	Avg: 1.9	Avg: 1.9
	Max: 2	Max: 2
Hausman	CHISQ(6) = 9.70	CHISQ(6) = 9.04
Test	p-value = 0.084	p-value = 0.108

Estimates with '*' ('**') indicate that the parameter is significantly different from zero at the five (one) percent level for a two-tailed test.

Table 4 shows the effects of explanatory variables on the revenue from fee-for-service per listed person and the total revenue from fee for service. *Patient Shortage* has a positive and statistically significant effect on the fee-for-service income per listed person, and a negative effect on the total income. Hence, we reject the hypothesis that more services to listed patients fully compensates for patient shortage. Also, from Table 4 neither *Gr-debt* nor *Gr-wealth* has an effect on service provision per listed person. However, there is a positive effect of these variables on the total fee-for-service income. Together these results imply that the additional income

comes from a larger patient list. Simultaneous doubling in *Gr-wealth* and in *Gr-debt* from the mean is predicted to increase fee-for-service income by 20 percent of the average annual fee-for-service income from national insurance among physicians in our sample.

Being a specialist in general medicine (*Gen-Med*) has a positive effect on both total and perpatient fee-for-service income. This is likely due to the fact that specialists in general medicine receive an additional fee per consultation from the national insurance. Also, from Table 4, a GP being male increases both the number of services per listed patient and the total fee-for-service income. The higher total income for male GPs is due to higher service intensity and longer lists.

We have also estimated the impact on preferred list size of gross wealth and gross debt by a regression model with random effects. Both variables are found to have a positive and statistically significant effect on preferred list size. Hence, this result supports prediction 4 of Section 3. We also find that being a specialist in general medicine, being male and being located in a municipality with a high level of centrality all contribute to a greater preferred list size. Being a specialist in community medicine and married both contribute to a small preferred list.

Municipality characteristics may be important determinants for physicians' decisions. A municipality's level of centrality picks up important location characteristics, and we have included three municipality dummies in the analyses. None turns out to be statistically significant in the analyses of hours of community services. In the analysis of fee-for-service income we only find that the highest level of centrality has a negative impact on the fee-for-service income per listed person. As an alternative to including dummies for municipality types we perform a two

level analysis of fee-for-service income with physicians nested in municipalities. These analyses do not change the sign and significance of the effects compared with the results in Table 4.

6. Concluding remarks

It is widely believed that many professionals hold high standards in how they should perform. Financial incentives are important, but not sufficient to determine their behaviors. Physicians are highly skilled professionals who have undertaken long trainings and maintain a commitment to the well-being of their patients. It is natural to expect that their behaviors are driven by a complex set of motives. In this paper, we have set out to investigate this set of motives for physicians in Norway.

We have shown that physicians respond to incentives in a heterogeneous way. Despite their lower remunerations, community services are undertaken by a significant fraction of physicians beyond the minimum required amount. We model this by postulating that GPs deriving utility from both financial returns and treating patients and performing tasks in the community health service. We proxy the preferences for community services with gross wealth and gross debt, and find them to be both statistical and quantitatively significant. Those GPs with lower gross wealth and gross debt tend to perform more community services; lower gross wealth and gross debt likely capture a more modest lifestyle and a stronger commitment to the service motive.

Policy implications of our study are important. Financial incentives cannot be expected to affect all physicians in a homogeneous way. Physicians likely respond to any set of incentives in

complex ways. In our study, lifestyles, proxied by physicians' gross wealth and gross debt, affect how they choose to supply community services. Much research is needed to identify other factors that contribute to their decisions.

Appendix: Physician deriving utility from serving patients in private practices

We now modify the utility function to check the robustness of results. We first let the utility function in (1) be modified to the following:

(1.A)
$$U(s,n,a) \equiv W(n,s) + \alpha sn + \beta a + \gamma n + \theta V(a) - C(ns+a).$$

Here the new term W(n,s) is the utility from providing care to n patients at the intensity of s services per patient. We assume that W is increasing and concave. We further specialize the function into two cases: (i) W takes the form nW(s), and (ii) W takes the form W(ns). Case (i) says that the physician derives a utility W(s) per private patient, and when there are n patients, the total utility is simply n times the per-patient utility. Case (ii) says that the physician derives a utility that is based on the aggregate services to all patients. Case (i) seems plausible, and we study it in some details. The analysis for Case (ii) is straightforward, and we will omit it.

We study the case when the quantity constraint $n \le D$ does not bind. The first-order condition with respect to n for the maximization of the modified utility function is

$$\left[\alpha s + W(s) + \gamma - sC'(ns + a)\right] = 0.$$

Dividing throughout by s, we get

135

$$\left[\alpha + \frac{W(s)}{s} - C'(ns + a)\right] = -\frac{\gamma}{s} < 0.$$

Next we consider the first-order derivative of the modified utility function with respect to s:

$$n[\alpha+W'(s)-C'(ns+a)]<0$$

where the inequality follows from the concavity of W (W'(s) < W(s)/s) and the preceding inequality (from the first-order condition with respect to n). Hence, the physician optimally chooses to lower the service per patient while choosing more patients.

In Case (i), the physician's altruistic preferences towards private patients is increasing in the services per patient, but at a decreasing rate. So a higher utility level may be achieved by simply adding more patients to the practice; more patients in the practice also mean more capitation income. For a general altruistic utility W(n,s), there may be a tendency for the service to rise above the minimum. This does not alter the fundamental incentive for increasing the patient list due to the capitation payment γ .

In a second variation of the utility modification, we can think of θ as a parameter that indicates a physician's tradeoff between monetary profit and private and community services. In this case, we modify the objective function accordingly:

(1.B)
$$U(s,n,a) = \alpha sn + \beta a + \gamma n + \theta [V(a) + W(s)] - C(ns + a)$$

Again the benevolent physician experiences some benefit from performing tasks in the community health service, V(a), and further experience some benefit from providing services in the private practice W(s). A physician having an objective function specified in (1.B) has an

altruistic attitude to providing services to the individuals who are actually listed in the practice, but this altruistic attitude is independent of list size.

We assume that W(s) is strictly concave and for simplicity we also assume that W(s) possesses properties that ensure that the physician chooses a service intensity in the interior of $[S_1, S_2]$. We study the case when the constraint $n \le D$ does not bind.

The first-order condition with respect to n for the maximization of (1.B) is

$$\left[\alpha s + \gamma - C'(ns + a)s\right] = 0$$

This can be expressed as:

$$\left[\alpha - C'(ns + a)\right]s = -\frac{\gamma}{s} < 0$$

Next we consider the first-order condition with respect to s:

$$[\theta W'(s) + \alpha n - C'(ns + a)n] = 0$$

This can be expressed:

$$\left[\alpha - C'(ns+a)\right]n = -\frac{\theta W'(s)}{n} < 0.$$

From these two first order conditions we get:

$$\frac{\theta W'(s)}{n^2} = \frac{\gamma}{s^2}.$$

The marginal benefit from service intensity is set proportional to the marginal benefit from the list size. In this version of the model, there is a tradeoff between service intensity and list size. Since the physician derives some utility from providing services in the private practice, he balances the incentive from the capitation payment γ from a longer list and low service intensity with the incentive to have high service intensity due to the service motives implicit in the function W(s).

By totally differentiating the system of equations implied by the three first-order conditions, we find that the comparative statics with respect to the altruism indicator are: $\frac{dn}{d\theta} < 0, \frac{ds}{d\theta} > 0 \text{ and } \frac{da}{d\theta} > 0.$ The results of the model specification implied by the objective function (1.B) are similar to those in section 3.

Acknowledgements:

We acknowledge financial support from the Research Council of Norway through the Health Economics Research Program at the University of Oslo (HERO). Data are provided by Norwegian Social Science Data Service (NSD). NSD is neither responsible for data analyses nor interpretations. For their comments, we thank conference and seminar participants at Boston University, European University Institute, University of Milan, the Workshop on Health Economics, Oslo 2006 and the Nordic Health Economics Meeting, Estonia 2007. We also thank Trond Olsen and Erik Biørn for their suggestions. The editor and two referees gave us very useful comments and suggestions.

References:

- Arrow, K. J., 1963. Uncertainty and the Welfare Economics of Medical Care. The American Economic Review 53, 941-973.
- Bazzoli, G. J., 1985. Does educational indebtedness affect physician specialty choice? Journal of Health Economics 4, 1-19.
- Cameron, A. C., Trivedi, P. K., 2005. Microeconometrics: Methods and applications Cambridge University Press, Cambridge.
- Carlsen, B., Norheim, O. F., 2003. Introduction of the Patient-list System in General Practice:

 Changes in Norwegian Physicians' Perception of their Gatekeeper Role. Scandinavian

 Journal of Primary Health Care. 21, 209-213.
- Culler, S.D., Bazzoli, G. J., 1985. The moonlighting decisions of resident physicians. Journal of Health Economics. 4, 283-292.
- Fox, M., 2003. Medical student indebtedness and the propensity to enter academic medicine. Health Economics. 12, 101–112.
- Godager G., Lurås H., 2005. I skyggen av Fastlegeordningen: Hvordan har det gått med det offentlige legearbeidet? (In Norwegian). HERO Working Paper, 2005:6 (Health Economics Research Programme at University of Oslo, Oslo).
- Godager G., Lurås H., 2006. Dual job holding GPs, the effect of patient shortage. Working paper.
- Gruber J., Owings, M., 1996. Physician financial incentives and cesarean section delivery. Rand Journal of Economics 27, 99-123.
- Iversen, T., 2005. A study of income-motivated behavior among general practitioners in the

- Norwegian list patient system. HERO Working Paper 2005:8 (Health Economics Research Programme at University of Oslo, Oslo).
- Iversen, T., Lurås, H., 2008. The impact of service provision on patient switching in a list patient system. HERO Working Paper 2008:4 (Health Economics Research Programme at University of Oslo, Oslo).
- Lurås, H., 2007. The association between patient shortage and patient satisfaction with general practitioners. Scandinavian Journal of Primary Health Care 25, 133-139...
- Moffit, R.A., 1999. New developments in econometric methods for labor market analysis. in:

 Orley C. Ashenfelter, David Card (Eds.), Handbook of Labor Economics, vol. 3. Elsevier,

 Amsterdam, pp. 1367-1397.
- Norwegian Official Statistics, 1999. Standard for regional classifications (in Norwegian language). Statistics. Norway, Oslo.
- Thornton, J., 2000. Physician choice of medical specialty: do economic incentives matter?

 Applied Economics. 32, 1419-1428.
- Tobin, J., 1958. Estimation of relationships for limited dependent variables. Econometrica. 26, 24-36.
- Yip, W. C., 1998. Physician response to Medicare fee reductions: changes in the volume of coronary artery bypass graft (CABG) surgeries in the Medicare and private sectors, Journal of Health Economics, 17, 675-699.