Journal of Health Economics 39 (2015) 159–170

Contents lists available at ScienceDirect

Journal of Health Economics

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

Competition, gatekeeping, and health care access

Geir Godager[a,*], Tor Iversen[a], Ching-to Albert Ma[b,c]

[a] Department of Health Management and Health Economics, University of Oslo, Norway
[b] Department of Economics, Boston University, USA
[c] University of Oslo, Norway

ARTICLE INFO

Article history:
Received 3 June 2013
Received in revised form
25 September 2014
Accepted 14 November 2014
Available online 22 November 2014

JEL classification:
D22
D43
H42
I10
I11

Keywords:
Primary care
Secondary care
Gatekeeping
Competition

ABSTRACT

We study gatekeeping physicians' referrals of patients to specialty care. We derive theoretical results when competition in the physician market intensifies. First, due to competitive pressure, physicians refer patients to specialty care more often. Second, physicians earn more by treating patients themselves, so refer patients to specialty care less often. We assess empirically the overall effect of competition with data from a 2008–2009 Norwegian survey, National Health Insurance Administration, and Statistics Norway. From the data we construct three measures of competition: the number of open primary physician practices with and without population adjustment, and the Herfindahl–Hirschman index. The empirical results suggest that competition has negligible or small positive effects on referrals overall. Our results do not support the policy claim that increasing the number of primary care physicians reduces secondary care.

1. Introduction

Many health care policy discussions are on primary care. In many European countries, each inhabitant must be enrolled with a primary care physician in order to receive national health services. In the United States, Title V of the Affordable Health Care Act provides subsidies for the training of primary care physicians and allied professionals (see http://www.healthcare.gov/law/full/index.html). Furthermore, Title IV of the Act promotes prevention, and it is expected that preventive care will be provided by primary care physicians.

Primary health care is less expensive than secondary and specialty care, so the emphasis on primary care for cost control is understandable. Perhaps, the most explicit cost-control perspective is the primary care physician's gatekeeping function. In many U.S. and European health plans, a patient can only obtain specialty care upon a referral made by her primary care doctor, also referred to as a gatekeeper. In this paper, we model the primary care physician's referral decision, and empirically assess the relationship between physician market conditions and gatekeeping.

A referral decision by a primary care physician or general practitioner (GP) likely depends on many factors such as medical conditions, current medical practice guidelines, availability of secondary care, the GP's service capacity, and financial incentives. The current policy recommendation of increasing the number of GPs adds one more dimension to the complex referral decision. Given a population of patients, more GPs will ultimately mean a more competitive market for doctors. This paper studies the relationship between competition in the GP market and a GP's referrals of patients to specialty care.

Such a study faces a number of difficulties. First, the number of GPs in any market changes slowly, even under any policy intervention. For example, subsidies in the U.S. Affordable Care Act are for physician training. This "natural" experiment will generate data only after many years, or perhaps even a decade. Similarly, in an experiment of a long duration, confounding factors affecting referral decisions will change over time. These changes may be difficult to track or be unobservable to the analyst. Second, in a multi-payer system such as the U.S., different health plans use different

* Corresponding author. Tel.: +47 22 84 50 29; fax: +47 22 84 50 91.
E-mail addresses: geir.godager@medisin.uio.no (G. Godager), tor.iversen@medisin.uio.no (T. Iversen), ma@bu.edu (C.-t.A. Ma).

http://dx.doi.org/10.1016/j.jhealeco.2014.11.005
0167-6296/

incentive contracts. Referral decisions likely will be influenced by these incentives. However, information on physician payment contracts is seldom available.

Our strategy is to use a cross-sectional data set, which can be interpreted as a snapshot that captures long-run changes, because different locations have had unique experiences for some time. (For example, differences between two countries at a given point in time result from long-term developments.) We use data from a 2008–2009 survey in Norway as our primary source, and supplement them with register data from Statistics Norway and from the National Health Insurance Administration. Because data are collected over a one-year period, time-varying confounding factors are irrelevant. However, our data include repeated municipality and GP observations, so we can account for unobserved variables at municipality and GP levels. The details of the survey and data are in the next section. Here, we would like to point out that (i) all self-employed Norwegian GPs are paid by the same financial contract, (ii) 95% of all Norwegians GPs are self-employed, and (iii) each Norwegian should be listed with a GP who is a gatekeeper for secondary care. In sum, problematic selection issues in multi-payer systems are avoided.

In Norway, a GP either lets his medical practice be open or closed to new patients. We use the number of open practices (with and without population density adjustments) as a measure of competition intensity in the GP market. The GP market is more competitive when there are more open practices because consumers have more options and each GP faces a more elastic demand. We also use the more conventional Herfindahl–Hirschman index as an alternative measure. Our use of the number of open practices in a geographical area as a measure of competition is uncommon.

Our empirical work seeks to explain specialty referral by competition intensity. We start with a model of GPs' referral decisions. As in much of the literature, we assume that a GP is guided by a profit motive and a concern for the patient. A GP's practice style is how he values a patient's potential benefits from specialty care and profits from providing services himself. Practice styles are assumed to be affected by market conditions. When the GP market becomes more competitive, the patient has more options. A GP who wants to retain a patient should adopt a practice style that values the patient's benefits. Competition may have a second effect. As the GP market becomes more competitive, a GP has less patients. Therefore, the GP incurs less disutility when he treats the patient himself. For a fixed reimbursement rate, the net profit from providing services increases.

Competition in the GP market has two opposing effects on referrals. More competition encourages a GP to show more concern for the patient, and therefore increases specialty referral. More competition also raises a GP's net profit for providing service himself, and therefore decreases specialty referral. Our model offers this new perspective, and we are able to assess empirically the overall effects of competition on specialty referrals.

The data sets allow us to control for patients' socioeconomic status, age and gender, as well as self-assessed health and chronic illness conditions. We also control for general and specialty health care access at the market level. Our (multinomial logit, and logit) regressions also account for clustering at the municipality levels. We find that competition either has insignificant or positive effects on GPs' referrals for patients to specialty care. In other words, we find no evidence that more competition among GPs will reduce their specialty referrals. Our results do not lend support to the secondary-care-reduction effect envisioned by a policy that promotes primary care.

Our data do not let us estimate separately the two opposing effects derived from our theoretical model. This, however, does not make our results less relevant. Our model of referral does capture the multi-faceted effects of competition on referrals, and an increase in primary care physician density results in more than a single change. This is an important aspect of the complexity in physician–patient interaction.

The literature on the primary and secondary health services is huge, whether that literature refers to health economics, health services research, or medicine. The health economics literature on the relationships between primary and secondary care is smaller but growing. In any case, the interest in primary care and health cost is topical. Using U.S. data, Baicker and Chandra (2004), and Chernew et al. (2009) find that the percentage of primary care physicians in a market is negatively associated with Medicare's reimbursement per beneficiary. Chernew et al. (2009), however, find no correlation between the percentage of primary care physicians and the growth in Medicare spending; thus Medicare policies that seek to reduce spending levels, but not growth rates, will ultimately fail to address cost issues.

Bradley Wright and Ricketts III (2010) use area-level data to show that within a location, a higher density of primary care physicians is associated with less inpatient admission and emergency room visits. Fortney et al. (2005) present results from a natural experiment at the U.S. Department for Veterans Affairs, in which the number of primary care facilities was increased in some districts but not in others. Using a difference-in-difference analysis of longitudinal data and instrumental variables for potential endogeneity problems, they find that an increase in primary care encounters is associated with a decrease in specialty medical encounters. Fortney et al. conclude that primary care is a substitute for specialist health care. Using survey data at the individual-patient level, Atella and Deb (2008) study whether primary care physicians and secondary specialists are substitutes or complements. They estimate a structural simultaneous-equation model where visits to different types of physicians are endogenous. When unobserved heterogeneity is appropriately accounted for, they find that primary care physicians and specialists are substitutes.

We model primary care physicians' referral decisions. The theoretical literature on referrals is quite rich. Barros and Olivella (2005) study cream skimming due to physicians in public services self-referring patients to their own private practices. Biglaiser and Ma (2007) examine the welfare effects of allowing dual practice and self-referrals. In our model the physician does not self-refer. Also, the referred specialists can reject referrals; existing papers in the literature have not considered this option.

Allard et al. (2011) consider how referral to secondary care is affected by incentive contracts for primary care physicians. Jelovac and Mariñoso (2003) compare optimal payment schemes with and without gatekeeping. Brekke et al. (2007) study the effect of GP gatekeeping on equilibrium quality in an imperfectly competitive secondary care market. González (2010) investigates the interactions between patients and GPs when some patients are informed about whether specialty care is appropriate. Our paper does not deal with the issues in these four papers. Our model is parsimonious, and focuses on competition in the GP market. Yet, it derives a set of predictions that we have taken to data.

The literature on competition in the health market is extensive; Gaynor and Town (2011) provide the latest review. It is fair to say that studies of competition have mainly focused on prices, qualities, costs, and health outcomes, and studies that use U.S. data outnumber those that use non-U.S. data. We are not aware of another paper that addresses the effect of competition in the primary care physician market on secondary care referral. Our paper therefore is the first to offer some evidence on this issue.

The common measures of competition in the literature are the number of providers (hospitals, physicians, nursing homes, etc.) within a geographical area, the n-firm concentration ratio, and the

Herfindahl–Hirschman index (see, for example, Wong et al. (2005)). In a patient-list system such as Norway, these measures do not capture the fact that patients can switch to another GP only if they can find open practices. We are unaware of any other study that uses the number of open GP practices as a measure of competition, with one exception: Iversen and Ma (2011) show that more intense competition, measured either by GP open practices or GP's desired practice sizes, significantly leads to more diagnostic radiology referrals.

The rest of the paper is organized as follows. Section 2 describes the Study Setting. A model of GP referral to either private or public specialist is set up in Section 3. Then we present our data set and descriptives in Section 4. The estimation results are in Section 5, and concluding remarks follow in Section 6. The Appendix contains a model for the competition between GPs, and the proofs of propositions.

2. Study setting

Norway has a three-tier government structure. At the top is the state, the next tier consists of 18 county councils, and the bottom tier consists of 430 municipalities. Norwegians' health care is covered by a decentralized national health services system. Municipalities are responsible for primary health care. Since 2002, secondary health care has been the responsibility of the state.

Each Norwegian is listed with a primary care physician, or a General Practitioner (GP). The Norwegian government is concerned about geographical distribution of GPs, and strictly regulates the entry of new GPs into municipalities. Licenses for GPs to practice at municipalities are allocated by the Directorate of Health. Patients may switch GPs twice a year, and in a year, about 3% of the patients do so. Patterns of patients switching physicians vary considerably, and depend on physician characteristics (Iversen and Lurås, 2011). Almost all GPs (95%) are self-employed. A GP typically contracts with his resident municipality.

The payment mechanism for a GP corresponds to a mixed system of a lump sum and fees. A GP's revenue can be divided into three roughly equal parts. First, he receives a capitation fee from the contracted municipality for each listed patient; this is the lump sum. Second, he receives fee-for-service reimbursements from the National Insurance Scheme (NIS, a public scheme that is an integral part of the state account) according to a fixed schedule negotiated between the state and the Norwegian Medical Association. Third, a GP receives copayments from patients for office consultations and tests. All fees and copayments are set at the national level, without any geographical variations. Hence, to GPs (and specialists) the payment system is entirely given. Each of these revenue sources accounts for about one third of a GP's practice income. As far as we can ascertain, GPs do not have financial arrangements with specialists; there are no side-contracts between them, and there is no evidence of any kickback from specialists due to GPs' referrals. (Moreover, GPs have no financial interests in laboratories.)

Almost all hospitals in Norway are public. Four Regional Health Authorities (RHAs) manage public hospitals. A number of private hospitals, both for-profit and not-for-profit, operate in Norway. Major private not-for-profit hospitals contract with the RHAs, and provide acute and elective care on the same terms as public hospitals. Private for-profit hospitals are very few and are geographically concentrated around Oslo, the capital. The RHAs buy some services from private for-profit hospitals. The remaining services are paid for by private health insurance or by patients directly.

Patients receive specialty outpatient consultations at public hospitals or at private offices. Specialists working at public hospitals receive salaries. Most private specialists contract with RHAs. Such a contract gives a private specialist an annual practice allowance from an RHA and fee-for-service reimbursements from the NIS. Private specialists are mainly located in urban areas. Approximately one third of all outpatient consultations are given by private specialists. A patient pays the same copayment whether the consultation is at a public hospital or a private office.

A GP is a gatekeeper. A GP must grant a referral before a patient receives specialty care at a public hospital or at a private office operated by a specialist under an RHA contract. A referral allows a patient to go to several visits for a defined medical condition within a year.

In 2009, the Norwegian government presented a major proposal for reforming the health care sector (White paper No. 47 (2008–2009)). The coordination reform was to solve problems due to low care quality for patients with chronic diseases, and high care costs. A part of the proposal suggested a substantial increase in GPs and related resources. However, due to lack of evidence, the Norwegian parliament rejected this part of the reform. Our research provides evidence for this policy discussion.

3. A model of referral

A patient is under the care of a primary care physician, or a General Practitioner (GP). The GP has to decide between treating the patient himself and referring the patient to secondary care. A referral can be made to a Public Specialist or a Private Specialist, who work, respectively, in the public and private sectors.

3.1. Patient, GP, and specialists

The patient is fully insured, and delegates treatment decisions to physicians. Let u denote the patient's benefit from the GP's treatment. This benefit depends on a patient's health status, and may take any value in an interval $[0, L]$. The GP observes this benefit u before making the referral decision. The GP does not know how much benefit the patient may obtain from secondary care, but believes that this is a random variable v. To simplify notation, we also let v vary on $[0, L]$, and it has a distribution F, and a density f. We assume that the distribution of v is independent of u. If they were correlated, we would simply replace the distribution F by a conditional one, but this would be conceptually similar.

Upon seeing the patient, a Specialist learns the value of v, as well as the value of u. Our interpretation is that the GP sends along the patient's medical information to the Specialist, who can infer the benefit u from primary care. There is a delay when a patient is referred to the Public Specialist, so the benefit becomes δv, $0 < \delta < 1$, if the patient is treated by the Public Specialist. Most public systems use waiting time as a rationing mechanism. This is true in Norway, and motivates our delay assumption. There is no delay when the patient is referred to the Private Specialist, so if the Private Specialist provides treatment, the patient's benefit is v.

A private physician working in the private sector is paid according to a fee-for-service contract with a national insurance system. The private GP has a fee-for-service rate p, while the Private Specialist has a rate q. We interpret p and q as unit profits, net of service costs. Very often we have $p < q$, so a Specialist receives a higher rate than a GP (although we do not use this inequality). The Public Specialist receives a salary. This difference in payments implies different incentives for service provisions.

Physicians behave as if their preferences are a weighted average of profits and patient's (expected) benefits. Physicians practice medicine according to professional protocols, but also care about profits. Alternatively, we can regard medical services as implicit or explicit bargaining outcomes between physicians and patients. Such outcomes (such as the Nash bargaining solution) are often a weighted average of a physician's profit and a patient's benefit. In the literature, a physician's trade-off between patient benefit and profit is said to constitute his practice style.

It is well known that physician practice styles vary widely. Definitive medical guidelines and protocols are not always available, and doctors operate under a gray area of medicine (Iversen and Lurås, 2000, and Chandra et al., 2011). It is well documented that the same medical condition may be treated differently across locations. Further, patients' preferences may have an effect on how physicians select treatments (Skinner, 2011). We further hypothesize that practice-style variations are influenced by GP market conditions. When GPs compete vigorously for patients, they may adjust their practice styles to put more weight on patient benefit. Alternatively, a more competitive GP market may endow patients with better outside options. This implies a better bargaining outcome for them.[1]

3.2. The referral process and physician utilities

We allow a Specialist the option to reject a referral and send the patient back to the GP for primary care services; this option is often ignored in the literature. The referral process is modeled as follows:

Stage 1: The GP observes the patient's benefit value u from his treatment, and decides between treating the patient, referring the patient to the Private Specialist, and referring the patient to the Public Specialist.

Stage 2: Upon a referral, the Private or Public Specialist gets to learn both u and v, the latter having being drawn according to distribution F. The Specialist decides between treating the patient and sending the patient back to the GP (who then has to treat the patient). There will be a delay if the referral has been to the Public Specialist, but there is no delay when the Specialist refers the patient back to the GP.

The GP's utility is $p+\alpha u$ if he treats the patient; he values profit p from fee-for-service (net) revenue and the patient's benefit at αu, where the strictly positive practice-style parameter α measures the importance of the patient's benefit. If the GP refers the patient to the Private Specialist, and the referral is accepted, his utility is αv. In this case, the GP no longer receives the fee-for-service payment p, but his concern for the patient remains. We continue to use the practice-style parameter α to measure how the GP values the patient's benefit v from specialty care. If the GP refers the patient to the Public Specialist, the patient experiences a delay, so the GP's utility is discounted by a factor of δ to $\alpha\delta v$. In the Norwegian system, the GP also receives a capitation payment for each patient under his care. At the time of referral, the GP has already received the capitation payment, so we do not write it down explicitly.

The Private Specialist's utility is similarly defined as $q+\beta v$ if he accepts the referral, where the practice-style parameter $\beta>0$ is the Specialist's weight on the patient's benefit. If he rejects the referral and sends the patient back to the GP, his utility is βu. Again, the Specialist values the patient's utility at the practice-style parameter β even when the patient is referred back to the GP.

The Public Specialist receives a fixed salary, so we let his payoff from treating the patient derive entirely from his concern for the patient. We normalize the Public Specialist's salary to 0, and his practice-style parameter to 1. When the Public Specialist sees the patient, the delay is already a sunk cost, so we write his utility from treating the patient as v, and his utility from sending the patient back to the GP as u.

3.3. Equilibrium decisions by specialists

The Public Specialist receives a salary and acts in the patient's best interest, so he treats the patient if $v>u$, and sends the patient back to the GP otherwise. The Private Specialist, however, may not act in the patient's best interest. If the Private Specialist accepts the referral, his payoff is $q+\beta v$. If he rejects the referral, the patient receives treatment from the GP, so the Private Specialist's payoff is βu. The Private Specialist accepts the referral if and only if $q+\beta v\geq\beta u$. Even when $v<u$, he may not redirect the patient back to the GP because he earns a monetary profit q by providing treatment. He will send the patient back to the GP only if the monetary payoff, q, is less than the incremental utility, $\beta(u-v)$, or equivalently $v<u-q/\beta$.

3.4. GP's utilities from treating and referring the patient

We now consider the GP's expected utilities from his three options in Stage 1. He anticipates both Specialists' best responses. First, if the GP treats the patient, his utility is

$$p+\alpha u. \tag{1}$$

Second, if the GP refers the patient to the Public Specialist, the referral will be accepted if and only if $v\geq u$. The GP's expected utility from this referral is

$$\int_{v<u}\delta[p+\alpha u]\mathrm{d}F(v)+\int_{v\geq u}\delta\alpha v\mathrm{d}F(v). \tag{2}$$

Here, the first integral (for $v<u$) corresponds to the Public Specialist rejecting the referral, so the GP's payoff is $p+\alpha u$, and this happens after a delay. The second integral (for $v\geq u$) corresponds to the Public Specialist accepting the referral, so the GP's payoff is αv, and again this happens after a delay.

Third, if the GP refers the patient to the Private Specialist, the referral will be accepted if and only if $q+\beta v\geq\beta u$. The GP's expected utility from referring the patient to the Private Specialist is

$$\int_{v<u-q/\beta}[p+\alpha u]\mathrm{d}F(v)+\int_{v\geq u-q/\beta}\alpha v\mathrm{d}F(v). \tag{3}$$

Here, the first integral (for $v<u-q/\beta$) corresponds to the Private Specialist rejecting the referral. The second integral (for $v\geq u-q/\beta$) corresponds to the Private Specialist accepting the referral.

3.5. GP's equilibrium referral

We assume that the parameter configuration in the model admits equilibria in which a referral may occur. The GP's equilibrium choice is obtained by comparing the treatment payoff $p+\alpha u$, and referral to Specialists' payoffs, respectively, expressions (2) and (3). His choice is guided by two considerations. First, the GP and the Private Specialist value both profits and the patient's benefit, but the Public Specialist is paid a salary, so is a perfect agent for the patient. Second, referring the patient to the Public Specialist means a delay. The first consideration is in favor of the GP referring the patient to the Public Specialist, but the second is against it.

First, suppose that u is large. The GP's best strategy is to provide treatment himself. There is only a small chance that a referral will

[1] Ellis and McGuire (1990), Fang and Rizzo (2009), and Chandra et al. (2011) show that provider competition may give patients more bargaining power.

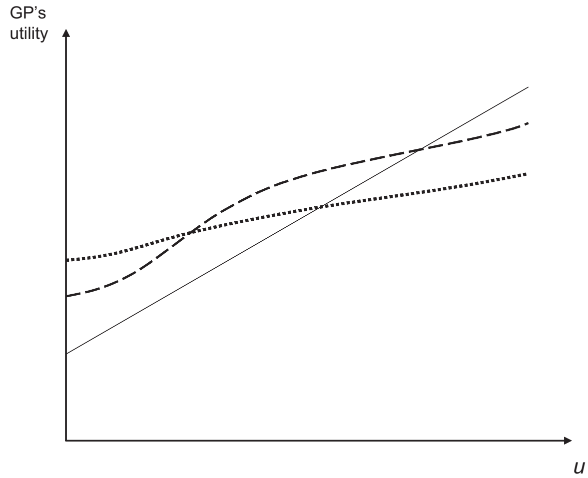

Fig. 1. GP's utilities from providing treatment and referrals.

show that v is higher than u. A referral to the Public Specialist is suboptimal because of the delay. A referral to the Private Specialist is suboptimal because the Private Specialist's concern for the patient is only partial. Second, suppose that u is small. Here, secondary care likely will benefit the patient. Referring the patient to the Public Specialist causes a delay, so the GP's equilibrium choice must be to refer the patient to the Private Specialist. Third, for medium values of u, the information about v is valuable, as in the second case. Here, the GP should also make a referral. However, whether the GP will refer the patient to the Public or Private Specialist depends on the discount factor δ, the parameters α and β, the fee-for-service rates p and q, and the distribution F.

Fig. 1 plots the GP's typical expected utilities from his three choices, as functions of the patient's benefit from the GP, u. The solid line is the GP's utility from treating the patient (expression (1)). The dashed line is the utility from referring the patient to the Public Specialist (expression (2)). The dotted line is the utility from referring the patient to the Private Specialist (expression (3)). The GP's equilibrium utility is given by the upper envelope of the three utility lines. In this example, at small u, the GP obtains the highest utility from referring the patient to the Private Specialist. For medium u, the highest utility is from referring the patient to the Public Specialist, while for large u, the highest utility is from the GP treating the patient himself. In general, however, each pair of the three expected utility lines can cross multiple times.

3.6. Competition and GP referrals

In the previous subsections, we have studied a GP's referral decision when he takes into account how the Specialists will respond. GPs compete against each other in the market. We can study a model of competition between GPs, and embed GPs' referral decisions in the previous subsections as subgames. In the Appendix, we have constructed a model of GP competition. Each GP is assumed to choose his practice style. Consumers have diverse preferences toward GPs (a horizontal dimension), but enjoy more from GP practice styles that give more weight to patient utilities (a vertical dimension). We use a Hotelling model and show that when competition becomes more intense, the equilibrium GP practice style will give more weights to patient utilities, so α increases.[2] The following propositions present how GPs' referrals vary when competition becomes more intense. (Proofs of Propositions are in the Appendix.)

Proposition 1. *In a more competitive market leading to a higher practice-style parameter α, the GP refers the patient to the Public Specialist or Private Specialist more often.*

A second way competition affects equilibrium referral is through its effect on the GP's net revenue p. If the total demand for GP services is constant, then as competition increases, each GP may have less patients. The GP has more leisure, so his disutility from work may decrease. This implies that the net revenue p may increase. The higher value of p implies a stronger incentive for the GP to provide treatment.

Proposition 2. *In a more competitive market leading to a higher GP fee-for-service rate p, the GP refers the patient to the Public Specialist or Private Specialist less often.*

A more competitive GP market means higher values of p and α. The above two propositions produce effects that act in opposite directions when competition becomes more intense. In fact, we can visualize these results from Fig. 1. An increase in α raises the slope of the solid line, (1), and both the slopes and intercepts of the other lines (2) and (3). The proof of Proposition 1 shows that the net effect will be an intersection of the solid line (1) with any of the other lines at a point further to the left. This means that the likelihood of referral increases. If the value of p goes up, the intercepts of all three lines in Fig. 1 rise. However, the intercept of the solid line (1) increases the most. As a result, the GP is more likely to keep his patient, so the likelihood of referral decreases.

Suppose that the GP has determined that keeping the patient is inferior, how is the GP's choice between referring the patient to the Public Specialist and the Private Specialist affected by competition? We have investigated the relative likelihood for these referrals, and the result is stated as follows.

Proposition 3. *A more competitive market has an ambiguous effect on the GP's choice between referring the patient to the Public Specialist and the Private Specialist.*

When a GP has determined that a patient is to be referred, the relative likelihood between a referral to a Public Specialist and to a Private Specialist may be increased or decreased when competition becomes more intense. This again reflects the multiple effects of competition on referral incentives. In Fig. 1, competition affects slopes and intercepts of the GP's utilities from referring the patient to the Public Specialist and Private Specialist, so their intersection point may move in either direction. In fact, the proof of Proposition 3 shows that each of a change in p and α has an ambiguous effect on the relative likelihood on referrals between the two specialists.

Our theoretical model has focused on referrals due to a GP's fee-for-service incentive. As we have said in Section 2, a GP also receives a capitation payment as well as patients' copayments. The capitation payment is a significant source of revenue, and a reason for a GP to want to retain patients in his list. The retention incentive may translate to a higher referral likelihood if patients think that health care provided by specialists is desirable. By being more eager to refer patients, a GP raises a patient's (expected) satisfaction level, and hence more likely retains the patient in the practice. Retention of patients is more important when a GP faces a lot of competition from other GPs, or equivalently, when patients have a lot of other GPs to choose from. The effect of more intense competition on the retention incentive is like a higher value of α, and it tends to increase GPs' referrals.

We have studied here the mixture of capitation and fee-for-service as a GP payment mechanism. Our theoretical model yields

[2] We suspect that other models of imperfect competition (such as logistic demand or even purely vertical differentiation) might yield the same result.

multifaceted effects on GP referrals to specialists when competition becomes more intense. The adaptation of our model to either a capitation or fee-for-service mechanism is straightforward, and the complexity of GP-competition effects on referrals remains true. Therefore, our insight continues to apply to other countries where payment mechanisms are similar.

4. Data and descriptives

Statistics Norway conducts an annual, cross-sectional "Survey of living conditions in Norway" (available at http://www.ssb.no/a/english/innrapportering/lev/). The main topic rotates. Every three or four years the population's self-assessed health and reported health care utilization will be the main topic. The main data for our study are from the Survey with health as the main topic conducted in 2008 and 2009 (Wilhelmsen, 2009). Statistics Norway drew a representative sample of 10,000 non-institutionalized residents aged 16 and above. The response rate was 66.8%. In total, 6465 face-to-face or telephone interviews were conducted. Compared with the Norwegian population, our sample is somewhat overly represented by women and those between 45 and 66 years old.

The Survey asked for information on living and health conditions such as common socioeconomic characteristics, self-assessed health status, etc. Data of respondents' income and education are obtained from the national registers. For our empirical work, the key information from the Survey includes the respondent's self-assessed health, number and types of chronic diseases, and use of primary and secondary health services during the twelve-month period before the survey. From the national registers, we obtain the identity of a respondent's regular GP. This information is merged with the survey data. We also add information of GPs and specialty care at the level of the respondent's municipality.

Our interest is to measure the intensity of competition between GPs. Most will agree that the total number of firms in a market captures competition intensity. Most also agree that in specific markets, the simplistic notion should be refined. Here, we attempt to accommodate specific details in the Norwegian GP market by three measures of competition intensity. First, we use the variable #OPEN; this is the number of GP practices in a respondent's municipality that accept new patients. Second, the variable #OPEN/CAPITA is #OPEN divided by the municipality's population measured in units of 10,000. Finally, the variable HERFINDAHL is the Herfindahl–Hirschmann index.

The variable #OPEN is perhaps the most straightforward. More GPs with open practices means more choices for those patients who are dissatisfied with their current GPs. A higher value of #OPEN, therefore, indicates a more competitive market for GPs.

Variable #OPEN, however, is rather simplistic. It ignores the distribution of open practices in municipalities. Consider two municipalities that have the same geographic size. However, one has twice the population than the other. Even with the same number of open GP practices, patients in the more densely populated municipality may have a smaller chance of finding a suitable GP. Now the variable #OPEN/CAPITA does capture the differences in population density in municipalities, so we use it as another measure of competition intensity.

Whereas #OPEN and #OPEN/CAPITA measure patients' potential choices when they consider switching GPs, the actual distribution of patients across different GPs may also measure the effect of competition. The Herfindahl–Hirschmann index is commonly used to describe market concentration. Its use for the health market is also common (Gaynor and Town, 2011, and Pauly, 2004). For a given market, the index is the sum of squares of each firm's market share; a firm's market share is the ratio of a firm's output to the total market output. For a monopoly, the index is 1, while for a market consisting of N firms each having the same market share, the index is $1/N\,(=\sum_{i=1}^{N}(1/N)^2)$. A smaller value of the index indicates a more competitive market.

In our case, we construct HERFINDAHL in the following way. For each GP, his output is the number of patients listed with him. A GP competes against other GPs in a market, which is defined as follows (see also Chen and Godager, 2011). For each GP in our data, we identify a geographical area within a 10-km radius from the center of his postal code. We call this a GP's circle. A GP competes against another GP if and only if their two circles intersect. The market for a GP consists of all the GPs with circles that overlap his. For each GP, we compute the Herfindahl index using patient lists of GPs in his market.

In our regressions, we seek to explain specialty care by market conditions. The three key variables to identify competition intensity are #OPEN, #OPEN/CAPITA, and HERFINDAHL. We do not use the list size of a respondent's GP because that is endogenous; physician supply may well correlate with patients' health status, and hence the specialty referral decision (see, for example, Dranove and Wehner, 1994). Our market-level measures for competition, however, are arguably exogenous or predetermined. Most important, our competition measures are aggregated over many physicians. A single GP's referral decision for a single patient cannot influence market competition intensity. Furthermore, whether a GP practice is open to new patients or not has been stable over time. Our regressions are identified by variations of #OPEN across municipalities, and variations of HERFINDAHL across markets. The #OPEN measure is based on municipality boundaries but the HERFINDAHL measure is not. Most Norwegians are listed with GPs in their home municipalities, but are not required to do so.

We use municipality-level indexes to control for patients' access to health care. These indexes are calculated in Lafkiri (2010), and are updates of those in Iversen and Kopperud (2005). We use two indexes: access to hospital care, and access to specialty care by private practitioners in a nonhospital setting. Access to hospital care is measured by the variable ACCESSPUB, which reflects hospital capacity in terms of the number of physicians, and is constructed as follows. Hospitals (in a municipality) have catchment municipalities. The variable ACCESSPUB is the number of physician specialists at public hospitals per 10,000 standardized inhabitants in the catchment municipalities. The standardization is according to automobile travel time between catchment municipalities and hospitals, with lower weights on populations farther away from hospitals. Access to private specialists, ACCESSPRIV, is constructed similarly, but now with the number of private specialists as the capacity measure. Indexes are standardized over the total number of Norwegian municipalities. Whereas the Herfindahl index is based on distances and the GPs' municipalities, all other capacity measures describe access in municipalities where patients reside. We do not have patients' travel distance information when they use secondary care.

Table 1 presents definitions of variables and descriptive statistics. In the sample, 83% of respondents reported seeing a GP in the twelve-month period before the interviews, while 40% reported having a consultation with a specialist. Since some patients visited both private and public specialists, the respective percentages of visits (at 20% and 27%, respectively) sum to more than 40%. On average, among respondents who had at least one GP consultation, they saw the GP more than 4 times. The corresponding average numbers of specialty consultations with the private specialist, public specialist, and any specialist are, respectively, 2.07, 2.46, and 2.70. Table 1 also presents respondents' gender and age information.

Table 1
Descriptives.

Variable	Definition	N	Mean	Std. Dev.	Min	Max
Variables at the municipality level						
#OPEN	No. of GPs who accept new patients	6416	37.8	66.23	0	213
#OPEN/CAPITA	No. GPs per 10,000 inhabitants who accept new patients	6416	3.89	2.72	0	24.3
HERFINDAHL	Herfindahl-index	6213	0.11	0.18	0	1
LOGHERFINDAHL	Ln(Herfindahl)	6213	−3.2	1.5	−5.96	0
ACCESSPRIV	Access indicator: private specialists	6425	0.7	1.12	−1.53	2.21
ACCESSPUB	Access indicator: public hospital physicians	6425	1.8	3.71	−1.59	11.78
Variables at the individual level						
CONSULT GP	Dummy = 1 if visited a GP during previous 12 months	6447	0.83	0.37	0	1
#CONSULT GP	No. of GP visits given the no. visits > 0	5311	4.16	5.53	1	130
CONSULT PRIV. SPEC.	Dummy = 1 if visited a private specialist during previous 12 months	6447	0.2	0.4	0	1
#CONSULT PRIV. SPEC.	No. of visits given the no. visits > 0	1279	2.07	3.95	1	100
CONSULT PUB. SPEC.	Dummy = 1 if visited a public specialist during previous 12 months	6447	0.27	0.44	0	1
#CONSULT PUB. SPEC.	No. of visits given the no. visits > 0	1726	2.46	3.5	1	55
CONSULT SPEC. CARE	Dummy = 1 if if visited specialist health service during previous 12 months	6447	0.4	0.49	0	1
#CONSULT SPEC. CARE	No. visits given the no. visits > 0	2553	2.7	4.29	1	102
MALE	Dummy = 1 if male	6447	0.49	0.5	0	1
YOUNG	Dummy = 1 if 16 years ≤ Age < 25 years	6447	0.14	0.35	0	1
MIDDLE1	Dummy = 1 if 25 years ≤ Age< 45 years	6447	0.34	0.47	0	1
MIDDLE2	Dummy = 1 if 45 years ≤ Age<67 years	6447	0.37	0.48	0	1
OLD	Dummy = 1 if 67 years ≤ Age	6447	0.15	0.35	0	1
VGOODHEALTH	Dummy = 1 if very good self-assessed health	6447	0.36	0.48	0	1
GOODHEALTH	Dummy = 1 if good self-assessed health	6447	0.44	0.5	0	1
CHRONIC	Dummy = 1 if at least one chronic disease	6435	0.4	0.49	0	1
INCOME	Gross household income in NOK 1000	6430	639	347	−1460	1460
HINC	Dummy = 1 if Income> Median of INCOME	6447	0.49	0.5	0	1
HIGHEDU	Dummy = 1 if > 14 years of education	6447	0.35	0.48	0	1

We use a number of variables to control for health status. Respondents were asked to rate their health in five grades. About 80% reported that their health status was either very good or good. The remaining 20% reported that their health status was fair, poor, or very poor. Although the self-assessed health status variables suggest a relatively healthy sample, 40% of the respondents had at least one chronic disease.

The mean (truncated) gross household income is 639,000 Norwegian Kroners (about US$106,000 at the approximate exchange rate of US$1 to 6Kr in 2008 and 2009). While there may be a nonlinear relationship between specialty-care utilization and a patient's income, we decided to exclude higher-order income terms to avoid multicollinearity. To capture this potentially nonlinear effect, in the estimations we use a binary variable HINC which is set to 1 for those with gross household income above the median, and 0 otherwise. Education may also have a nonlinear effect. There may be significant utilization differences between those with higher education and those without. Again, we use a binary variable HIGHEDU which is set to 1 for those with 14 or more years of education. In the sample, 35% of the respondents achieved at least two years of education beyond the high school.

On average, 37.8 GPs in the respondent's municipality would accept new patients. This corresponds to 45% of all GPs in the sample. Adjusting for the population size, we find a mean of 3.89 for #OPEN/CAPITA,which is the number of open GP practices per 10,000 of inhabitants. The variable of HERFINDAHL has a mean of 0.11. The means of the two standardized variables of health care access ACCESSPUB and ACCESSPRIV are 1.80 and 0.70, respectively.

5. Estimation and results

Our model describes a GP choosing between providing service to a patient himself, referring the patient to a Private Specialist, and referring the patient to a Public Specialist. Propositions 1–3 characterize the GP's decision as the competition intensity changes. Generally, competition has ambiguous effects, so our goal is to estimate the overall impact of competition intensity on GPs' secondary care referrals. Our estimation strategy also takes into account a hierarchical structure in which patients are clustered in GPs' lists, and GPs are clustered in municipalities. In our estimation sample, the number of patients listed with the same GP varies between 1 and 22, with a median of 2. The number of individuals residing in the same municipality varies between 1 and 742, with a median of 30. Even though our data do not contain time varying variables, our multilevel data allow us to account for unobserved variables, both at the municipality and GP levels, by exploiting the repeated observations of municipalities and GPs.

The theoretical model naturally suggests a discrete-choice empirical specification, so we use a multilevel multinomial logit regression. The endogenous variable has three categories: a patient using a Private Specialist, using a Public Specialist, and the status quo of using a GP for treatment. To account for unobserved heterogeneity at the physician and municipality levels, the model includes two random intercepts corresponding to these two levels. Three models are fitted, each corresponding to one of three competition-intensity measures of the GP market. We estimate using GLLAMM in STATA 13 (Rabe-Hesketh et al., 2005), and present the results with robust standard errors in Table 2. Because there are three discrete choices, we present the estimated effects of independent variables in terms of relative risk ratios. For logit regressions, the relative risk ratios are independent of the explanatory variables, so are convenient.[3]

[3] Let p_{ij} be the probability that individual i chooses alternative j, and let m index be the base category. In the multinomial logit, the relative risk ratio of individual i choosing alternative j is $\frac{p_{ij}}{p_{im}} = \exp(\beta_j' \boldsymbol{x}_i)$ (see, for example, Cameron and Trivedi, 2005, p500). The log of the relative risk ratio is linear: $\ln\left(\frac{p_{ij}}{p_{im}}\right) = \beta_{j0} + \beta_{j1}x_{i1} + \ldots + \beta_{jk}x_{ik} + \ldots + \beta_{jK}x_{iK}$. The value of β_{jk} measures how the log of the relative risk ratio changes when the regressor x_{ik} increases by 1. We have $\ln\left(\frac{p_{i'j}/p_{i'm}}{p_{ij}/p_{im}}\right) = \beta_{jk}$, where $\ln\left(\frac{p_{i'j}}{p_{i'm}}\right) = \beta_{j0} + \beta_{j1}x_{i'1} + \ldots + \beta_{jk}x_{i'k} + \ldots + \beta_{jK}x_{i'K}$, with $\boldsymbol{x_{i'}} = \boldsymbol{x_i}$, except $x_{i'k} = x_{ik} + 1$. The value of β_{jk} is therefore the log relative risk ratio for alternative j versus m when x_{ik} increases by one unit. The relative risk ratio of the regressor x_{ik}

Table 2
Utilization of Private Specialist or Public Specialist relative to GP service. Multilevel multinomial logit regressions with random effects at physician and municipality levels.

	Model 1				Model 3	
	Private Specialist Relative risk ratio (SE[a])	Public Specialist Relative risk ratio (SE[a])	Private Specialist Relative risk ratio (SE[a])	Public Relative risk ratio (SE[a])	Private Relative risk ratio (SE[a])	Public Relative risk ratio (SE[a])
#OPEN	1.004 (0.003)	1.001 (0.003)				
#OPEN/CAPITA			0.972 (0.020)	1.025 (0.015)		
LOG HERFINDAHL					0.908* (0.045)	0.988 (0.041)
ACCESSPRIV	1.174** (0.054)	0.924 (0.038)	1.173** (0.056)	0.928 (0.039)	1.128 (0.056)	0.924 (0.041)
ACCESSPUB	0.941 (0.050)	0.984 (0.048)	1.010 (0.014)	1.009 (0.014)	0.995 (0.014)	1.007 (0.013)
MALE	0.643** (0.056)	0.716** (0.048)	0.643** (0.056)	0.718** (0.047)	0.657** (0.056)	0.719** (0.048)
YOUNG	0.735 (0.120)	0.868 (0.090)	0.734 (0.120)	0.865 (0.089)	0.747 (0.120)	0.872 (0.091)
MIDDLE2	1.256** (0.108)	1.198 (0.123)	1.258** (0.108)	1.192 (0.122)	1.262** (0.109)	1.194 (0.124)
OLD	1.494** (0.185)	1.199 (0.143)	1.508** (0.185)	1.186 (0.142)	1.519** (0.192)	1.196 (0.146)
VGOODHEALTH	0.804 (0.108)	0.358** (0.040)	0.800 (0.106)	0.358** (0.040)	0.814 (0.108)	0.364** (0.041)
GOODHEALTH	0.868 (0.103)	0.551** (0.054)	0.868 (0.102)	0.551** (0.054)	0.899 (0.106)	0.566** (0.057)
CHRONIC	1.518** (0.120)	1.977** (0.174)	1.512** (0.119)	1.981** (0.175)	1.526** (0.118)	2.018** (0.173)
HINC	1.274* (0.138)	1.094 (0.074)	1.268* (0.137)	1.100 (0.074)	1.293* (0.149)	1.089 (0.075)
HIGHEDU	1.201* (0.101)	1.327** (0.089)	1.208* (0.101)	1.331** (0.089)	1.188 (0.105)	1.328** (0.089)
MALE_GP	0.966 (0.093)	0.972 (0.081)	0.958 (0.089)	0.971 (0.081)	0.974 (0.093)	0.972 (0.081)
MIDDLE1_GP	1.206 (0.119)	1.047 (0.066)	1.205 (0.118)	1.039 (0.066)	1.190 (0.121)	1.035 (0.066)
CONSTANT	0.201** (0.032)	0.953 (0.147)	0.231** (0.040)	0.866 (0.152)	0.154** (0.032)	0.906 (0.162)
# OF OBS	5827		5827		5725	
# OF PHYSICIANS	2548		2548		2500	
# OF MUNICIPALITIES	169		169		167	
LOG LIKELIHOOD	−5504.6353		−5501.1898		−5401.7185	

Estimates with *(**) indicate that the relative risk ratio is significantly different from 1 at the 5% (1%) level for a two-tailed test.
[a] Reported standard errors are robust.

The competition-intensity measures #OPEN and #OPEN/CAPITA have no significant effect on the probability of GPs referring a patient to specialty care at private or public settings; the estimated relative risk ratios for #OPEN and #OPEN/CAPITA are not statistically significantly different from 1. The third competition-intensity measure, LOGHERFINDAHL, has no effect on referrals to Public Specialists, but has a negative and significant effect (at 5%) on Private Specialist referrals. A higher value of LOGHERFINDAHL means a more concentrated, hence less competitive, market, so the estimated relative risk ratio below 1 implies that GPs refer patients to Private Specialists less often as the GP market becomes less competitive. (Conversely, GPs refer patients to Private Specialists more often as the market becomes more competitive.) Indeed, the Herfindahl index increasing by 1% reduces the relative risk of a patient being referred to a Private Specialist by 9%. Based on our results, we have found no evidence that a more competitive GP market, as indicated by any of our three measures, will lead to a reduction in speciality care referral.

Our results cannot be attributed GPs' tendency to locate in municipalities where illnesses (and hence referrals) are more prevalent. This is because our market variables, #OPEN and #OPEN/CAPITA, measure excess capacities. We do not use GPs' actual list sizes except in the construction of the Herfindahl–Hirschman index. There is no reason to expect a positive correlation between excess capacity and illness. (In fact, a negative correlation is more likely because GPs have more cases to take care of when they are located in areas with sicker patients.)

The reverse association that more specialty referrals result in more open GP practices, is implausible. An open practice implies excess capacity, which does not carry any financial reward under the current payment system. By contrast, a GP receives a capitation payment when he enrolls a patient. If the GP opens his practice to new patients, the capitation payment must be attractive to him. More specialty referrals mean less treatment provisions, so a GP loses the fee-for-service payment, too.

In all regressions, we have controlled for gender, age, socioeconomic status and self-assessed health. The estimated effects of controls are as expected. Being female raises the probability of a referral to specialty care at both private and public settings. Senior citizens are more likely to have a referral than the young, but the effects are only significant for referrals to Private Specialists. Individuals with chronic diseases have a higher probability of being referred, and this applies to both Private and Public Specialists. Individuals' self-assessed health influences the probability of referral to Public Specialists but not to Private Specialists. However, the reverse is true for individuals' incomes. Interestingly, well-educated individuals have higher probabilities of using specialty care. Finally, we use ACCESSPRIV and ACCESSPUB to control for access to private and public specialty care. Although ACCESSPRIV has the expected positive and significant effects on referrals to Private Specialists, ACCESSPUB does not.

5.1. Robustness checks

For robustness checks, we consider two other estimations. First, we use logit models to estimate separately referrals to Private and Public Specialists for each of the three competition-intensity measures. These logit models focus on a GP's referrals to a particular secondary care setting. The results from these six regressions are reported in Table 3. Second, we fit logit models to estimate the probability of referrals to either Private or Public Specialists and present the results in Table 4. In these two tables, choices are assumed to be binary, and the presented relative risk ratios therefore coincide with odds ratios.

In the first three columns of Table 3, the dependent variable takes the value 1 when the patient is referred to a Private Specialist, and 0 otherwise. In the last three columns of Table 3, the dependent variable takes the value 1 when the patient is referred to a Public Specialist, and 0 otherwise. Results of the logit models in

for alternative j is given by $\exp(\beta_{jk})$. Hence, a relative risk equal to 1 means that β_{jk} is 0.

Table 3
Logit regressions on utilization of Private and Public Specialists. Population averaged panel data models with random effects at municipality level.

	Private Specialist			Public Specialist		
	Relative risk ratio (SE[a])	Relative risk ratio (SE[a])	Relative risk ratio (SE[a])	Relative risk ratio (SE[a])	Relative risk ratio (SE[a])	Relative risk ratio (SE[a])
#OPEN	1.001 (0.001)			1.002 (0.002)		
#OPEN/CAPITA		0.990 (0.014)			1.021 (0.011)	
LOGHERFINDAHL			0.941* (0.029)			1.040 (0.027)
ACCESSPRIV	1.200** (0.043)	1.217** (0.040)	1.182** (0.044)			
ACCESSPUB				0.961 (0.034)	0.987 (0.009)	0.994 (0.009)
MALE	0.577** (0.034)	0.575** (0.034)	0.579** (0.034)	0.763** (0.040)	0.763** (0.040)	0.758** (0.040)
YOUNG	0.690** (0.084)	0.687** (0.085)	0.697** (0.084)	0.941 (0.083)	0.937 (0.083)	0.942 (0.086)
MIDDLE2	1.124 (0.094)	1.126 (0.095)	1.126 (0.095)	1.150 (0.091)	1.147 (0.092)	1.149 (0.094)
OLD	1.425** (0.157)	1.432** (0.160)	1.431** (0.160)	1.137 (0.100)	1.131 (0.101)	1.134 (0.103)
VGOODHEALTH	0.644** (0.065)	0.642** (0.065)	0.647** (0.063)	0.376** (0.036)	0.375** (0.036)	0.380** (0.038)
GOODHEALTH	0.833* (0.073)	0.834* (0.074)	0.848 (0.073)	0.575** (0.047)	0.575** (0.048)	0.584** (0.051)
CHRONIC	1.511** (0.102)	1.513** (0.103)	1.516** (0.101)	1.944** (0.129)	1.953** (0.130)	2.006** (0.133)
HINC	1.191* (0.091)	1.192* (0.093)	1.209* (0.096)	1.032 (0.058)	1.035 (0.059)	1.017 (0.060)
HIGHEDU	1.229** (0.080)	1.235** (0.082)	1.203** (0.082)	1.184** (0.067)	1.189** (0.068)	1.201** (0.070)
MALE_GP	1.072 (0.077)	1.071 (0.077)	1.072 (0.079)	0.969 (0.062)	0.971 (0.062)	0.973 (0.061)
MIDDLE1_GP	1.074 (0.087)	1.080 (0.087)	1.076 (0.091)	0.985 (0.051)	0.980 (0.052)	0.967 (0.049)
CONST	0.207** (0.031)	0.217** (0.036)	0.176** (0.032)	0.508** (0.063)	0.468** (0.066)	0.551** (0.074)
# OF OBS	6315	6315	6201	6315	6404	6201
No. MUNICIPALITIES	174	174	172	174	176	172
OBS.PER MUNICIPALITY						
MIN	1	1	1	1	1	1
AVG	36	36	36	36	36	36
MAX	742	742	738	742	742	738

Estimates with *(**) indicate that the relative risk ratio is significantly different from 1 at the 5% (1%) level for a two-tailed test.
[a] Reported standard errors are robust.

Table 3 resemble those in the multinomial logit models in Table 2, both in terms of which regressors being significant and their quantitative effects. The variables #OPEN and #OPEN/CAPITA have no effect on the probability of GPs referring a patient to specialty care at private or public settings. The third competition-intensity measure, LOGHERFINDAHL, has no significant effect on referrals to Public Specialists, but has a negative and significant effect (at 5%) on referrals to Private Specialists. The negative effect says that as competition becomes more intense, GPs refer patients to Private Specialists more often.

In the models presented in Table 4, the dependent variable takes the value 1 when the patient is referred to a specialist, irrespective of whether the referral happens at a public or private setting, and takes the value 0 otherwise. Now the variables #OPEN,

Table 4
Logit regressions on utilization of Specialist, either Private or Public. Population averaged panel data models with random effects at municipality level.

	Specialist		
	Relative risk ratio (SE[a])	Relative risk ratio (SE[a])	Relative risk ratio (SE[a])
#OPEN	1.002 (0.002)		
#OPEN/CAPITA		1.004 (0.012)	
LOGHERFINDAHL			0.978 (0.031)
ACCESSPRIV	1.020 (0.031)	1.020 (0.033)	1.018 (0.033)
ACCESSPUB	0.975 (0.032)	1.005 (0.007)	1.000 (0.010)
MALE	0.642** (0.035)	0.642** (0.035)	0.644** (0.035)
YOUNG	0.840* (0.075)	0.839* (0.075)	0.847 (0.075)
MIDDLE2	1.234** (0.088)	1.232** (0.087)	1.235** (0.089)
OLD	1.388** (0.133)	1.386** (0.132)	1.393** (0.136)
VGOODHEALTH	0.443** (0.040)	0.443** (0.040)	0.450** (0.040)
GOODHEALTH	0.622** (0.052)	0.622** (0.052)	0.638** (0.054)
CHRONIC	2.000** (0.117)	1.999** (0.118)	2.043** (0.118)
HINC	1.145* (0.068)	1.147* (0.069)	1.139* (0.071)
HIGHEDU	1.204** (0.067)	1.210** (0.067)	1.206** (0.068)
MALE_GP	0.984 (0.064)	0.981 (0.063)	0.989 (0.063)
MIDDLE1_GP	1.046 (0.065)	1.043 (0.066)	1.031 (0.067)
CONST	0.769* (0.101)	0.767 (0.109)	0.709* (0.110)
# OF OBS	6315	6315	6201
No. MUNICIPALITIES	174	174	172
OBS PER MUNICIPALITY			
MIN	1	1	1
AVG	36	36	36
MAX	742	742	738

Estimates with *(**) indicate that the relative risk ratio is significantly different from 1 at the 5% (1%) level for a two-tailed test.
[a] Reported standard errors are robust.

#OPEN/CAPITA and LOGHERFINDAHL, have no significant effect on the probability of GPs referring a patient to specialty care. The estimated effects of our controls are as expected and resemble those in Tables 2 and 3.[4]

6. Concluding remarks

In many policy discussions, the idea of having more primary care physicians has been promoted. It is thought that this will lead to better primary care, and reduce secondary care at the same time. We have constructed a model of GPs' secondary and specialty care referrals. Then we have used a set of Norwegian survey data and the corresponding register data to test whether competition among GPs will lead to more or less referrals.

We model the referral decision of a GP who is paid by both capitation and fee-for-service. Our theory predicts two opposing effects when the GP market becomes more competitive. First, GPs become more concerned about patients' welfare as the GP market becomes more competitive, so they refer patients to secondary care more often. Second, competition may reduce GPs' workload, so they earn a higher net profit from providing treatments to patients themselves, and refer patients to secondary care less often.

Using data from a representative survey of Norwegian citizens conducted by Statistics Norway in 2008–2009 and linked data of survey respondents' GPs and municipalities, we assess the overall effects of competition. We find no evidence that more competition in the GP market will reduce speciality care. Our three competition measures, #OPEN, #OPEN/CAPITA, HERFINDAHL, either have insignificant effects or small and positive effects on the likelihood of a referral and the number of specialty visits.

Both #OPEN and #OPEN/CAPITA consider numbers of GPs with open lists but not the total number of open places. For a patient who considers switching GPs, what counts is whether or not a GP accepts new patients rather than the number of patients that he accepts. As a check of our argument, we have run all regressions with the number of open places per capita in a municipality as the competition indicator. We have not found statistically significant effects (and have not presented them here).

Can our results be due to GPs with lower qualities having open practices? Consumers actively seek out better GPs, so lower-quality GPs tend to refer their patients more often for specialty care in order to keep them. This mechanism may well be at work at the level of the individual GP. However, our competition indicators are at the municipality level. The "quality" argument would then require a concentration of low-quality physicians within municipalities with a lot of GPs. We cannot disregard this possibility although we are unaware of any evidence supporting such a claim.

How relevant are our results to other settings? The Norwegian system has some distinct features. Referrals at public and private settings have the same copayment. Private Specialists receive fee-for-service, whereas Public Specialists are on salaries. The GP is paid a combination of fee-for-service and a fixed capitation fee. Patients' free choice of GPs and GPs' gatekeeping under a capitation fee are main drivers behind our results of competition. Therefore, our results are of interest to any health system in which patients choose GPs, and GPs act as gatekeepers and rely on both capitation and fee-for-service revenues. Among the European countries these characteristics apply to Denmark, the Netherlands, and Italy according to Paris et al. (2010). With the recent introduction of free choice of GP in the UK, our results are also relevant to there.

The basis for a policy to increase primary care physicians seems straightforward. When there are more GPs, patients have more choices, and likely receive more care from them. Nevertheless, the relationship between GP competition and secondary care is multifaceted. One might even argue that making referrals is one of the GP's responsibility, so increasing the number of GPs will increase referrals.

A policy that aims to reduce costly specialty care seems to require a change in medical practices. Success in cost reduction is achieved when secondary care is substituted by primary care. This means that the traditional guidelines for GPs' and specialists' responsibilities have to be redrawn. Alternatively, an integrated approach, in which GPs and specialists together internalize cost and benefit, may offer a better avenue for efficiency.

Acknowledgement

We are grateful to the Research Council of Norway for financial support through the Health Economics Research Program at the University of Oslo (HERO). Tinna Asgeirsdottir, and seminar participants at Boston University, Indiana University Purdue University Indianapolis, Stony Brook University, Universidad Computense de Madrid, the 31st Nordic Health Economists Study Group Meeting in Umeå, Sweden, the 4th Biennial Conference of the American Society of Health Economists in Minneapolis, the 8th European Conference on Health Economics in Helsinki, and the 8th iHEA World Congress in Toronto gave us useful comments. We thank an editor and three reviewers for their valuable suggestions.

Appendix

GP competition

We now embed the GP's equilibrium decisions into a model of imperfect competition between GPs. Suppose that a GP has a practice style α. Given this, his equilibrium treatment and referral decisions can be derived. These decisions yield continuation payoffs for a patient. Consider now the *ex ante* stage when a patient's utilities u and v, respectively, from treatment by a GP and a Specialist are unknown. Let $V(\alpha)$ denote a patient's expected utility when he is listed a GP with practice style α. We let V be increasing and concave; that is, a patient gets a higher (expected) utility from a GP whose practice style gives more weight to patient utility, but this is diminishing.

It is simplest to consider a Hotelling model. Suppose that one GP is located at one end of a line of unit length. Consumers are uniformly distributed on the line. In Stage 1, each profit-maximizing GP chooses his practice style to compete for patients; in Stage 2, each consumer chooses a GP. The choice of a practice style is similar to offering an overall care quality to potential patients. Suppose that the GP at location 0 sets his practice style at α_1 and the other GP sets his practice style at α_2. A consumer located at x, $0 \le x \le 1$, receives expected utilities $V(\alpha_1) - tx$ and $V(\alpha_2) - t(1-x)$, respectively, from listing with these GPs. Given these practice styles, the shares of consumers listing with the GP at 0 and the GP at 1 are, respectively,

$$\left[\frac{1}{2} + \frac{V(\alpha_1) - V(\alpha_2)}{2t}\right] \quad \text{and} \quad \left[\frac{1}{2} + \frac{V(\alpha_2) - V(\alpha_1)}{2t}\right].$$

A GP gains a higher market share by being more generous.

Let P be the profit if the GP does not refer a listed patient. This expected profit under nonreferral P is to be distinguished from the

[4] We also considered modeling a GP's decision as a multistage process. Here, the GP first decides between providing treatment and referring the patient to specialty care. Then he chooses between a Private Specialist and a Public Specialist if he decides not to provide treatment himself. However, our data do not allow us to identify this system.

fee-for-service net revenue rate p because P includes the capitation fee. A practice style α commits a GP to make referral decisions as if his preferences were $p+\alpha u$, so there will be referrals. This implies that the GP will not gain the profit P under a practice style $\alpha>0$. The reduction from P is denoted by $c(\alpha)$, which is assumed to be increasing and convex. The two GPs' profits are, respectively,

$$[P-c(\alpha_1)]\left[\frac{1}{2}+\frac{V(\alpha_1)-V(\alpha_2)}{2t}\right] \quad \text{and}$$
$$[P-c(\alpha_2)]\left[\frac{1}{2}+\frac{V(\alpha_2)-V(\alpha_1)}{2t}\right].$$

In an equilibrium each GP chooses his practice style to maximize profit. The first-order condition for profit maximization for the GP at location 0 is

$$[P-c(\alpha_1)]\frac{V'(\alpha_1)}{2t}-c'(\alpha_1)\left[\frac{1}{2}+\frac{V(\alpha_1)-V(\alpha_2)}{2t}\right]=0.$$

We consider a symmetric equilibrium. The equilibrium practice style α is obtained by setting $\alpha_1=\alpha_2$ in the above first-order condition:

$$P-c(\alpha)=\frac{c'(\alpha)}{V'(\alpha)}t.$$

We totally differentiate the above to obtain:

$$-c'(\alpha)\mathrm{d}\alpha=\frac{c'(\alpha)}{V'(\alpha)}\mathrm{d}t+\left[\frac{V'(\alpha)c''(\alpha)-V''(\alpha)c'(\alpha)}{V'(\alpha)^2}\right]\mathrm{d}\alpha.$$

This simplifies to

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t}=-\frac{[(c'(\alpha))/(V'(\alpha))]}{c'(\alpha)+[(V'(\alpha)c''(\alpha)-V''(\alpha)c'(\alpha))/(V'(\alpha)^2)]}<0,$$

where the inequality follows because each term inside the square brackets is positive from the convexity of c and the concavity of V. A reduction in t is usually interpreted as a more competitive market in the Hotelling model. Therefore, we conclude that as the market becomes more competitive, the equilibrium practice style becomes more generous.

Proof of Proposition 1: A GP refers a patient if and only if his expected utility from referral is higher than the utility from treating the patient. This is equivalent to

$$p+\alpha u\le\min\left\{\begin{array}{l}\displaystyle\int_{v<u}\delta[p+\alpha u]\mathrm{d}F(v)+\int_{v\ge u}\delta\alpha v\mathrm{d}F(v),\\ \displaystyle\int_{v<u-q/\beta}[p+\alpha u]\mathrm{d}F(v)+\int_{v\ge u-q/\beta}\alpha v\mathrm{d}F(v)\end{array}\right\},\quad(4)$$

the two terms on the right-hand side of the inequality being the expected utilities from referrals to the Public and Private Specialists, respectively.

The derivative of (1) with respect to α is u. The derivative of (2) with respect to α is

$$\int_{v<u}\delta u\mathrm{d}F(v)+\int_{v\ge u}\delta v\mathrm{d}F(v).\quad(5)$$

Finally, the derivative of (3) with respect to α is

$$\int_{v<u-q/\beta}u\mathrm{d}F(v)+\int_{v\ge u-q/\beta}v\mathrm{d}F(v).\quad(6)$$

Suppose that at some u, say $\widehat{u}$, we have

$$p+\alpha\widehat{u}=\int_{v<\widehat{u}}\delta[p+\alpha\widehat{u}]\mathrm{d}F(v)+\int_{v\ge\widehat{u}}\delta\alpha v\mathrm{d}F(v)$$

so that the GP is indifferent between treating the patient and referring the patient to the Public Specialist. We rewrite this condition as

$$p+\alpha\widehat{u}=\delta pF(\widehat{u})+\alpha\left\{\int_{v<\widehat{u}}\delta\widehat{u}\mathrm{d}F(v)+\int_{v\ge\widehat{u}}\delta v\mathrm{d}F(v)\right\}.\quad(7)$$

Because $p>\delta pF(\widehat{u})$, we have

$$\alpha\widehat{u}<\alpha\left\{\int_{v<\widehat{u}}\delta\widehat{u}\mathrm{d}F(v)+\int_{v\ge\widehat{u}}\delta v\mathrm{d}F(v)\right\}.$$

This inequality says that at $\widehat{u}$, the derivative of (1) with respect to α, namely $\widehat{u}$, is strictly smaller than (5), the derivative of (2) with respect to α.

At $\widehat{u}$, as α increases, the GP's expected utility from referring the patient to the Public Specialist increases faster than the utility from treating the patient. Because of (7), we conclude that at $\widehat{u}$ the GP refers the patient to the Public Specialist when α increases.

Next, suppose that at some u, say $\widetilde{u}$, we have

$$p+\alpha\widetilde{u}=\int_{v<\widetilde{u}-q/\beta}[p+\alpha\widetilde{u}]\mathrm{d}F(v)+\int_{v\ge\widetilde{u}-q/\beta}\alpha v\mathrm{d}F(v)$$

so that the GP is indifferent between treating the patient and referring the patient to the Private Specialist. We rewrite this condition as

$$p+\alpha\widetilde{u}=pF(\widetilde{u}-q/\beta)+\alpha\left\{\int_{v<\widetilde{u}-q/\beta}\widetilde{u}\mathrm{d}F(v)+\int_{v\ge\widetilde{u}-q/\beta}v\mathrm{d}F(v)\right\}.\quad(8)$$

Because $p>pF(\widetilde{u}-q/\beta)$, we have

$$\alpha\widetilde{u}<\alpha\left\{\int_{v<\widetilde{u}-q/\beta}\widetilde{u}\mathrm{d}F(v)+\int_{v\ge\widetilde{u}-q/\beta}v\mathrm{d}F(v)\right\}.$$

This inequality says that at $\widetilde{u}$, the derivative of (1) with respect to α, namely $\widetilde{u}$, is strictly smaller than (6), the derivative of (3) with respect to α.

At $\widetilde{u}$, as α increases, the GP's expected utility from referring the patient to the Private Specialist increases faster than the utility from treating the patient. Because of (8), we conclude that at $\widetilde{u}$, the GP refers the patient to the Private Specialist when α increases. □

Proof of Proposition 2: With respect to p, the derivative of the GP's utility from treating the patient in (1) is 1. The derivative of the GP's utility from referring the patient to the Public Specialist (2) is

$$\int_{v<u}\delta\mathrm{d}F(v)<1.$$

Finally, the derivative of the GP's utility from referring the patient to the Private Specialist (3) is

$$\int_{v<u-q/\beta}\mathrm{d}F(v)<1.$$

Clearly, the GP's utility rises faster in p when he treats the patient than when he refers the patient to either the Public or Private Specialists.

If the GP treats a patient at some u, we have

$$p+\alpha u\ge\max\left\{\begin{array}{l}\displaystyle\int_{v<u}\delta[p+\alpha u]\mathrm{d}F(v)+\int_{v\ge u}\delta\alpha v\mathrm{d}F(v),\\ \displaystyle\int_{v<u-q/\beta}[p+\alpha u]\mathrm{d}F(v)+\int_{v\ge u-q/\beta}\alpha v\mathrm{d}F(v)\end{array}\right\}.\quad(9)$$

As p increases, the utility term on the left-hand side of (9) increases at a rate of 1, but each utility term on the right-hand side of (9) increases at a rate less than 1. Hence, the inequality in (9) remains valid as p increases. □

Proof of Proposition 3: The GP's utilities from referring the patient to the Public Specialist and the Private Specialist are, respectively, in (2) and (3). Suppose that at some u they are equal:

$$\int_{v<u} \delta[p+\alpha u]\mathrm{d}F(v) + \int_{v\geq u} \delta\alpha v \mathrm{d}F(v)$$
$$= \int_{v<u-q/\beta} [p+\alpha u]\mathrm{d}F(v) + \int_{v\geq u-q/\beta} \alpha v \mathrm{d}F(v).$$

Because

$$\int_{v\geq u-q/\beta} \alpha v \mathrm{d}F(v) > \int_{v\geq u} \delta\alpha v \mathrm{d}F(v) \tag{10}$$

we have

$$\int_{v<u} \delta[p+\alpha u]\mathrm{d}F(v) > \int_{v<u-q/\beta} [p+\alpha u]\mathrm{d}F(v). \tag{11}$$

The derivatives of (2) and (3) with respect to α are, respectively (see also (5) and (6)),

$$\int_{v<u} \delta u \mathrm{d}F(v) + \int_{v\geq u} \delta v \mathrm{d}F(v) \text{ and } \int_{v<u-q/\beta} u \mathrm{d}F(v) + \int_{v\geq u-q/\beta} v \mathrm{d}F(v).$$

Now inequalities (10) and (11) imply

$$\int_{v\geq u-q/\beta} v \mathrm{d}F(v) > \int_{v\geq u} \delta v \mathrm{d}F(v) \text{ and } \int_{v<u} \delta u \mathrm{d}F(v) > \int_{v<u-q/\beta} u \mathrm{d}F(v).$$

Therefore, the derivative of (2) may be bigger or smaller than the derivative of (3). We conclude that when α increases, the GP's utility from referring the patient to the Public Specialist may increase more or less than from referring the patient to the Private Specialist.

Next, consider the derivatives of (2) and (3) with respect to p. These are, respectively:

$$\int_{v<u} \delta \mathrm{d}F(v) \text{ and } \int_{v<u-q/\beta} \mathrm{d}F(v).$$

Again, the relative magnitude between these two derivatives is undetermined. We conclude that when p increases, the GP's utility from referring the patient to the Public Specialist may decrease more or less than from referring the patient to the Private Specialist. □

References

Allard, M., Jelovac, I., Léger, P.T., 2011. Treatment and referral decisions under different physician payment mechanisms. Journal of Health Economics 30, 880–893.

Atella, V., Deb, P., 2008. Are primary care physicians, public and private sector specialists substitutes or complements? Evidence from a simultaneous equations model for count data. Journal of Health Economics 27, 770–785.

Baicker, K., Chandra, A., 2004. The productivity of physician specialization: evidence from the Medicare program. American Economic Review 94, 357–361.

Barros, P.P., Olivella, P., 2005. Waiting lists and patient selection. Journal of Economics & Management Strategy 14, 623–646.

Biglaiser, G., Ma, C.A., 2007. Moonlighting: public service and private practice. RAND Journal of Economics 38, 1113–1133.

Bradley Wright, T., Ricketts III, C., 2010. The road to efficiency? Re-examining the impact of the primary care physician workforce on health care utilization rates. Social Science & Medicine 70, 2006–2010.

Brekke, K.R., Nuscheler, R., Straume, O.R., 2007. Gatekeeping in health care. Journal of Health Economics 26, 149–170.

Cameron, A.C., Trivedi, P.K., 2005. Microeconometrics: Methods and Applications. Cambridge University Press, Cambridge, UK.

Chandra, A., Cutler, D., Song, Z., 2011. Who ordered that? The economics of treatment choices in medical care. In: Pauly, M.V., McGuire, T.G., Barros, P.P. (Eds.), In: Handbook of Health Economics, vol. 2. Elsevier Science, Amsterdam, pp. 397–432.

Chen, X., Godager, G., 2011. Development of Competition Indicators in the Norwegian General Practice: Constructing a Postal Code-Specific Herfindahl–Hirschman Index Applying STATA Software. Health Economics Research Program at the University of Oslo, Norway, HERO Working Paper 2011:1.

Chernew, M.E., Sabik, L., Chandra, A., Newhouse, J.P., 2009. Would having more primary care doctors cut health spending growth? Health Affairs 28, 1327–1335.

Dranove, D., Wehner, P., 1994. Physician-induced demand for childbirths. Journal of Health Economics 13, 61–73.

Ellis, R.P., McGuire, T.G., 1990. Optimal payment systems for health services. Journal of Health Economics 9, 375–396.

Fang, H., Rizzo, J.A., 2009. Competition and physician-enabled demand: the role of managed care. Journal of Economic Behavior & Organization 72, 463–474.

Fortney, J.C., Steffick, D.E., Burgess Jr., J.F., Maciejewski, M.L., Petersen, L.A., 2005. Are primary care services a substitute or complement for specialty and inpatient services? Health Services Research 40, 1422–1442.

Gaynor, M., Town, R.J., 2011. Competition in health care markets. In: Pauly, M.V., Mcguire, T.G., Barros, P.P. (Eds.), In: Handbook of Health Economics, vol. 2. Elsevier Science, Amsterdam, pp. 499–637.

González, P., 2010. Gatekeeping versus direct-access when patient information matters. Health Economics 19, 730–754.

Iversen, T., Kopperud, G.S., 2005. Regulation versus practice: the impact of accessibility on the use of specialist health care in Norway. Health Economics 14, 1231–1238.

Iversen, T., Lurås, H., 2000. Economic motives and professional norms: the case of general medical practice. Journal of Economic Behavior & Organization 43, 447–470.

Iversen, T., Lurås, H., 2011. Patient switching in general practice. Journal of Health Economics 30, 894–903.

Iversen, T., Ma, C.A., 2011. Market conditions and general practitioners' referrals. International Journal of Health Care Finance and Economics 11, 245–265.

Jelovac, I., Mariñoso, B.G., 2003. Gps' payment contracts and their referral practice. Journal of Health Economics 22, 617–635.

Lafkiri, K., 2010. Municipality Level Accessibility to Specialized Health Care in Norway. Health Economics Research Program at the University of Oslo, Norway, HERO Working Paper 2010:7.

Paris, V., Devaux, M., Wei, L., 2010. Health Systems Institutional Characteristics: A Survey of 29 OECD Countries. OECD Publishing, http://dx.doi.org/10.1787/5kmfxfq9qbnr-en, OECD Health Working Papers, No. 50.

Pauly, M.V., 2004. Competition in medical services and the quality of care: concepts and history. International Journal of Health Care Finance and Economics 4, 113–130.

Rabe-Hesketh, S., Skrondal, A., Pickles, A., 2005. Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. Journal of Econometrics 128 (2), 301–323.

Skinner, J., 2011. Causes and consequences of regional variations in health care. In: Pauly, M.V., Mcguire, T.G., Barros, P.P. (Eds.), In: Handbook of Health Economics, vol. 2. Elsevier Science, Amsterdam, pp. 45–93.

White paper No. 47, 2008–2009 to the Storting: The Coordination Reform: Proper Treatment – at the Right Place and Right Time. The Government of Norway, Oslo.

Wilhelmsen, M., 2009. Samordnet levekårsundersøkelse 2008 - Tverrsnittsundersøkelsen. Dokumentasjonsrapport. Statistisk sentralbyrå, Oslo.

Wong, H.S., Zhan, C., Mutter, R., 2005. Do different measures of competition measures matter in empirical investigations of hospital behavior? Review of Industrial Organization 26, 61–87.