GATEKEEPING – OPEN DOOR TO EFFECTIVE MEDICAL CARE UTILISATION?

Eva Hromádková

CERGE-EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

WORKING PAPER SERIES (ISSN 1211-3298)
Electronic Version

400

Working Paper Series **400**
(ISSN 1211-3298)

Gatekeeping – Open Door to Effective Medical Care Utilisation?

Eva Hromádková

CERGE-EI
Prague, November 2009

EVROPSKÝ SOCIÁLNÍ FOND
PRAHA & EU: INVESTUJEME DO VAŠÍ BUDOUCNOSTI

Projekt je financován Evropským sociálním fondem, rozpočtem hl. města Prahy a státním rozpočtem

ISBN 978-80-7343-202-7 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-191-3 (Národohospodářský ústav AV ČR, v.v.i.)

Gatekeeping - Open Door to Effective Medical Care Utilisation?*

Eva Hromádková†
CERGE–EI‡

Abstract

We assess the ability of health insurance plans with gatekeeping restrictions to control the utilization of medical care through their influence on the choice of the initial provider. Empirical results are based on the individual-level utilization panel data from 2001-2006 Medical Expenditure Panel Survey. We find only small differences between the initial provider chosen by individuals enrolled in gatekeeping and non-gatekeeping plans. This, together with the fact that within gatekeeping plans, 21 percent of patients self-refer to specialists, imply that the intended cost-containment effect of gatekeeping, namely restricting the utilization of specialty care, is surprisingly weak.

Abstrakt

V naší práci hodnotíme efektivitu plánů zdravotního pojištění v USA za použití tzv. "gatekeeping" restrikcí (t.j. s restrikcí primárního poskytovatele zdravotní péče na vybraného všeobecného lékaře) při regulaci poptávky a následné spotřeby zdravotní péče. Naše zjištění jsou založena na analýze dat z Výběrového panelového šetření výdajů na zdravotní péči - Medical Expenditures Panel Survey, provedeném na reprezentativním vzorku americké populace v letech 2001-2006. Výsledky analýzy ukazují, že při výběru poskytovatele prvotního kontaktu s zdravotní péčí je mezi pojištěnci s a bez "gatekeeping" restrikcí jen malý rozdíl, což je překvapující zejména vzhledem k explicitnímu zaměření této restrikce. Taktéž jsme zjistili, že skoro 21 procent pacientů s "gatekeeping" restrikcí, kteří by jako prvního měli kontaktovat svého všeobecného lékaře a žádat jeho doporučení, navštíví přímo specialistu. Z těchto zjištění vyplývá, že zamýšlený efekt této restrikce, tedy regulace využívání specialistů, je překvapivě slabý.

Keywords: health insurance, gatekeeping, health care utilization, episodic demand model, initial provider
JEL classification: I11, I19

*The author would like to thank Randall K. Filer, Partha Deb, Mike Grossman, Ted Joyce, Frantisek Kopriva and Jan Hanousek for helpful comments. The views expressed are those of the author and do not necessarily reflect the position of any of the affiliated institutions.

†Email: Eva.Hromadkova@cerge-ei.cz

‡A joint workplace of the Center for Economic Research and Graduate Education, Charles University, Prague, and the Economics Institute of the Academy of Sciences of the Czech Republic. Address: CERGE–EI, P.O. Box 882, Politických vězňů 7, Prague 1, 111 21, Czech Republic

1 Introduction and Motivation

Many health insurance companies in the US try to contain costs by influencing the behavior of enrollees on the specifics of how they demand medical care. This linking of health insurance with medical care provision frequently takes the form of managed care and is implemented as a complex set of requirements and incentives for both patients and physicians. One rarely challenged cornerstone of managed care is gatekeeping - the requirement that a primary care physician (PCP)[1] coordinates a patient's treatment and provides referrals to specialists, hospitals or other medical care. The intended purpose of gatekeeping is to ration access to more costly specialty care only to patients who really need such care, and retain treatment of less severely ill patients with the less costly PCP. In this paper, we use panel data to evaluate whether gatekeeping actually works as intended and to shed some light on patterns of subsequent care by examining patient and physician behavior in gatekeeping and non-gatekeeping situations.

As health care costs continue to escalate, it has become progressively more imperative for health insurance companies to find effective ways to control costs while not significantly degrading patient outcomes. To this end, a large body of empirical literature has attempted to quantify the effects of managed care (Miller and Luft, 1997 and 2002) and, more specifically, gatekeeping (Pati, 2005; Deb and Trivedi, 2006a) on medical care utilization. The results so far have been inconclusive.

This lack of conclusive findings arises from two methodological problems, both of which this paper addresses. First, previous work routinely uses aggregate data on total utilization of medical care in a given time period. Looking at the end point, i.e., total use, of an often multi-layered care process tells us little about the

[1]By primary care physician we mean family doctors and general practitioners, who provide primary care services for patients in both gatekeeping and non-gatekeeping programs.

effect of gatekeeping, which by design, should manifest itself most strongly at the beginning by influencing who the patient chooses as their initial contact provider (ICP).[2] Second, previous studies fail to address the problem of self-selection, i.e., the fact that enrollees who prefer direct access to the specialty care because of their anticipated health care needs should prefer to opt for plans without gatekeeping restrictions.

To address the end point/aggregate data issue, we use detailed, individual-level panel data from the 2001-2006 Medical Expenditure Panel Survey, in which we can distinguish not only each episode of illness, but within a given episode, the pattern of initial and subsequent provider contact. To our knowledge, use of this design is unique in the literature. To deal with the self-selection problem, we take advantage of a quasi-natural experiment within our data by identifying the impact of gatekeeping in a sub-sample of respondents whose enrollment choice between gatekeeping and non-gatekeeping insurance plans was driven by the factors external to their health care demand - the choice of insurance plans provided by their employer.

Our main finding is intuitively surprising and has profound implications for health insurance design. Contrary to previous studies, we find no economically significant difference between the initial provider contact (IPC) choices of individuals enrolled in gatekeeping or non-gatekeeping plans. Even though one would predict fewer self-referrals to specialists, and consequently an increase in the number of PCP visits for patients in gatekeeping as opposed to non-gatekeeping plans, our results show only slight differences that are partially explained by self-selection. Therefore, we conclude that gatekeeping does not effectively influence patient behavior through its presumed main operating channel, i.e., restricting direct access to specialty care. The implication of these findings for the insurance industry is

[2]Initial contact provider categories as defined in this paper are: PCP, specialist, non-MD, hospital, and emergency room.

that gatekeeping has a minimal impact on health care costs.

While the above finding is provocative, it is not the entire gatekeeping story. What if gatekeeping through other mechanisms, such as patient and physician incentives, influences the specifics of whether a patient utilizes primary care or the broad spectrum of specialist care accessible after the ICP? We address this issue by analyzing patterns of medical care after initial contact for both gatekeeping and non-gatekeeping plans. We find that contact with different types of ICP imply distinctly different patterns of future care (e.g. utilization rates after initial hospitalization are much higher than after initial contact with PCP). However, conditional on the same type of ICP, we find no differences between the utilization measures of gatekeeper versus non-gatekeeper patients, with the following important exception. Gatekeeper patients are 25 percent more likely to be referred by their PCP to a specialist than non-gatekeeper patients who use PCP as ICP. This result continues to support the main finding that gatekeeping does not reduce specialty care.

Following these hints of unexpected interactions within the gatekeeping model, we delve further into the nuances of gatekeeping's impact on patient and physician behavior. We analyze the effectiveness and appropriateness of referrals by PCPs as compared to self-referrals, with the focus on gatekeeping plans. We find that PCP-referred patients have greater illness severity (as proxied by the probability of further hospitalization) than self-referred patients. Therefore, in keeping with expectations, it appears that within gatekeeping plans, PCPs do retain the healthier patients, referring only the more severely ill patients to specialty care. On the other hand, a sizeable 21 percent of patients with gatekeeping do self-refer mainly for conditions that are recognizable and less severe, but still suitable for specialty care even though they may have to pay out of pocket. This efficiency shift in the composition of self- and PCP- referred patients within gatekeeping plans needs to be taken into account by policy makers, especially when combining gatekeeping

with referral incentives for physicians.

The remainder of the paper is organized as follows. Section 2 presents a theoretical discussion of the link between insurance status and utilization as well as empirical work preceding this study. Section 3 describes the conceptual model of decision-making in the demand for health care; section 4 presents the construction and characteristics of the data 4; section 5 outlines the methodology; the empirical results are discussed in section 6; and section 7 presents the conclusions drawn from this study.

2 Managed care, gatekeeping and health care utilization

2.1 Theoretical approaches

The idea behind managed care (MC) is to design measures that affect both the demand and supply side of health care system in order to limit increasing medical care utilization and, consequently, rising health care expenditures. If one aims to analyze how the design of these measures specifically affects the demand side, the key requirement is to understand the decision-making process underlying the health care demand at the level of an individual. Two broad categories of models attempt to tackle this problem. One line of reasoning, initiated by the seminal model of Grossman (1972) views the demand for health care solely as an outcome of the agent's own utility maximization, where health enters the agent's utility as valuable capital, and demand for health care is derived the same way as the demand for any other investment. The other line of thought, represented by Zweifel's (1981) principal-agent approach assumes that demand for health care is determined by the physician who, due to an informational advantage, acts as an effective agent for patient.

The model that best fits a general understanding of the decision-making process underlying health care demand is a combination of the above mentioned approaches. Based on the episodic model of care developed by Stoddart and Bauer (1981) and applied by Pohlmeier and Ulrich (1995) or Holmes and Deb (1998), one can model this process as having two parts with each part better explained by one of the models. In the first stage, it is the patient who decides whether to visit a physician at all and, specifically, what type of provider he wants to contact first. This decision can be satisfactorily captured by a Grossman-type model because it is based on the patient's optimization with respect to his budget constraints and supply-side restrictions. On the other hand, after first contact, the patient for practical purposes delegates the decision about the future course of treatment to the physician, who then, in line with Zweigel's model, determines further medical care utilization based on both clinical reasons and his own financial incentives[3]. Various mechanisms of managed care are targeted to affect both of these stages. In our paper, we focus particularly on gatekeeping, which at the first stage restricts provider choice, and then affects the further course of treatment through referral policies.

Gatekeeping in health care is a system where the PCP coordinates patient care and provides referrals to specialists, hospitals and other medical services. The standard rationale for introducing gatekeeping into MC is based on moral hazard in its standard interpretation by Arrow (1963) or Pauly (1968). According to moral hazard theory , people with insurance (i.e. people who do not face the real price of the provided health services at the time they use them) tend to demand more services than they otherwise would without insurance. The gatekeeping role of the PCP in this context would lie in rationing access to specialty care only to patients who need it to reduce unnecessary medical interventions, thus controling costs.

[3]For empirical evidence see e.g. Campbell XXX.(2007).

This cost-containment function of gatekeeping is one of two fundamental arguments put forward by Scott (2000), the other being the informational advantage of the PCP over the patient. Naturally, one would expect the PCP to be better informed about the quality and appropriate specialization of secondary care for a particular patient. On the other hand, a PCPs actions in terms of diagnosis, treatment and/or referral strongly affect the patient's welfare, but the patient is not fully aware of how much influence the PCP has or whether the PCP's action is appropriate in the particular situation. Therefore, the value of gatekeeping depends on the validity of the assumption that primary care can effectively substitute for more costly specialty and inpatient care.

From the theoretical point of view, however, the complexity of interaction between primary and secondary care makes this assumption look oversimplified. Fortney, Steffick, Burgess Jr., Maciewski and Petersen,(2005) provide an overview of mechanisms that can lead to both substitutability and complementarity of primary and secondary care. Primary care can be seen as a substitute for secondary care if (1) it averts the need for specialist care by the prevention or early detection of illness; (2) by managing chronic illnesses, it avoids their evolution into more severe cases treatable only at higher levels of expertise, or (3) by simply restricting the access through formal rules such as gatekeeping (Starfield, 1994).[4] On the other hand, several possible mechanisms exist by which primary care acts as a complement to secondary care: by using services that are ancillary to primary care (like laboratory tests) or by detection of illnesses that have to be treated through secondary care (like cancer). Therefore, we cannot predict the overall effect of gatekeeping as a policy aimed at increasing use of primary care and we have to rely on analysis of real world examples to provide comprehensive insight into the interactions with other regulatory mechanisms and, eventually, on effect on the behavior of patients.

[4]In their analysis of a mixed public-private health care system in Italy, Atella and Deb (2008) found that PCP, public specialists, and private specialista are indeed substitute sources of care.

2.2 Empirical findings

This paper contributes to a broad class of empirical investigation on the effects of managed care on the utilization of medical services. Comprehensive reviews of this line of research are provided in Glied (2000), and Miller and Luft (1997, 2002). As Glied points out, however, the concept of managed care incorporates many different combinations of the mechanisms used to manage health care provision and utilization. By using the broad categorization health management organization (HMO) versus "other" health care plans, many studies fail to take this aspect into account. So, it is not surprising that very little evidence exists on specifically how gate-keeping and corresponding provider access restrictions affect medical care utilization. Also, the problem of self-selection into health insurance plans is of a great importance to this line of research. Naturally, people who prefer unrestricted access to specialist care because of their anticipated health needs should opt for plans without gatekeeping restrictions. [5] Therefore, if the authors of a particular study do not account for selection, they easily overestimate the effect of gatekeeping by capturing the effect of unobserved characteristics of enrollees rather than that of gatekeeping alone.

In the literature, researchers employ different techniques to deal with the problem of selection into particular insurance types. First, they avoid the issue altogether by taking advantage of various natural or randomized experiments and consecutive random assignment of enrollees. Martin, Diehr, Price and Richardson (1989) use a randomized trial to determine the effectiveness of a gatekeeping plan that imposes cost-containment incentives both on PCP (fundholding) and patient (100 percent copayment for self-referral) and find that it reduces the costs of ambulatory services by reducing specialist visits. For the second widely used

[5] Glied (2002) claims that the results of the studies on selection (see e.g. Hellinger, 1995) suggest that managed care plans have 20-30 percent prior utilization advantage over indemnity plans.

technique, instrumental variable estimation, finding variables that are both good predictors of insurance choice and exogenous to further medical care utilization is difficult. Moreover, although this approach works well when outcomes of interest can be modeled using linear regression methods, it is difficult to apply for utilization measures because of their count data character.[6]

Another option is to use latent factor models that parametrically account for selection on unobservables and then estimate the model using simulation based methods. This approach is taken in Deb and Trivedi (2006a) and represents one of the few studies specifically examining the effect of gatekeeping and the physician network attributes of the health insurance plans on health care utilization. The authors define health insurance plans as bundles of three possible restrictions: (1) providers' network; (2) sign-up with PCPs, and; (3) out-of network costs coverage and then estimate their effect on five fairly general measures of health care utilization. The results of the study indicate significant evidence of selection into managed care plans [7]. The authors find that if the plan requires a PCP sign-up, individuals have more contacts with non-physician providers and also undergo significantly more surgeries and hospital stays, a finding that the authors explain by a tendency of PCPs to diagnose new medical conditions requiring further treatment within the procedure of routine examinations.

The second line of research relevant to my topic is literature on the channels through which gatekeeping restrictions work. The first channel is through imposing constraints and incentives on physician referral behavior. Grembowski, XXX (1998) model the expected influence of managed care on physician referrals and health outcomes. Similarly to the previously cited Fortney, XXX (2005), they point out that by reducing or delaying access to specialist services, MC can affect health

[6] See Mullahy (1997) for the nonlinear instrumental variable methods for count data models.

[7] In their complementary paper, Deb and Trivedi (2006b) claim that this selection is based on socio-economic observables rather than health status.

outcomes both in a positive or negative way. [8] Forrest, Nutting, Starfield and von Schrader(2002) present a descriptive overview of family physicians' referral decisions finding that apart from clinical reasons, one of the common determinants of referral is patient pressure. Also in one-third of cases, the referral was made during encounters other than office visits, which clearly undermines the cost-containment function of gatekeeping.

The other aim of gatekeeping is to restrict the widespread practice of self-referral as it has become a common route to specialty care. The analysis of the NAMCS survey of office-based physicians by Forrest and Reid (1997) found that 31 percent of specialists' new patients were self-referred. In attempting to evaluate the appropriateness of these self-referrals, the authors used hospitalization as a proxy for the severity of the illness and found that self-referred patients have a lower probability of hospitalization than patients referred by a physician, and therefore the appropriateness of self-referrals can be questioned.

A theoretically important tool to discipline patient behavior in order to contain costs is changing the out-of-pocket price of treatment whereby under gatekeeping, patients usually have higher co-payments or must even bear the full cost of the medical procedure if they decide to bypass the gatekeeper. Holmes and Deb (1998) examine the ways in which the costs of nonresidential mental health care depend on the choice of the initial provider and the level of cost sharing imposed on the patient. The results are consistent with an episodic model of demand. The out-of-pocket price significantly influences the patient's choice of the initial provider, but the utilization after that appears to be unrelated to the financial incentives that patients face. Pati, Shea, Rabinowitz and Carrasquillo (2005) use MEPS data to

[8] On the one hand, most medical problems can be diagnosed and treated appropriately in primary care (Donaldson, M.D., et al., 1996). As more patients are allocated to PCPs to diagnose and treat, however, this can result in worse health outcomes due to misdiagnosis, inappropriate delay of referrals, or simply the provision of lower quality of care than specialists would deliver (Kassirer, 1994).

look at the effect of managed care gatekeeping on overall health care costs. They find that mean per capita expenditures were approximately 6 percent lower for gatekeeping plan enrollees compared to indemnity plan enrollees, primarily due to lower out-of-pocket expenditures.

3 Methodology

We base our methodology on the episodic model of health care demand. We implement it in three steps corresponding to the three main stages prior to and during the medical care episode: choice of the insurance plan, initial contact, and further course of treatment. The main questions we then address are: (1) Does gatekeeping affect the choice of the initial contact provider? (2) Does this effect translate into further medical utilization?

3.1 Choice of insurance plan

Before the actual utilization, an individual decides whether and what kind of health insurance to purchase. In the US, generally, one is eligible for a publicly provided insurance (e.g. Medicare, Medicaid, or SCHIP); can purchase private insurance through one's employer, insurance group or on the individual market; or can remain uninsured. In our study we focus on a more subtle decision concerning the type of private insurance, specifically one with and without gatekeeping restrictions. People make their decision by comparing their options in terms of offered health insurance products with their needs, i.e. expected health care utilization, and financial constraints. While supply-side characteristics like the availability of insurance through one's employer or the variability of insurance products offered are exogenous to the choice of the insurance type, the expected health care needs are not and, therefore, they are a source of self-selection and endogenenity, which

we have to deal with in the estimation of the second stage.

3.2 Choice of initial provider

After a person becomes ill and seeks medical help, he has to decide what type of physician or medical care provider he will contact first. This choice is influenced by his personal characteristics, previous experience (knowledge of doctors, expectations about quality of treatment), an subjective evaluation of the severity of the illness, and the conditions and incentives embedded in his health insurance plan. This is the stage of the decision-making targeted to be influenced by gatekeeping, and where we attempt to estimate the role of gatekeeping on the choice of ICP.

We use two techniques to deal with the problem of endogeneity. The first, propensity score matching, is based on the idea that we can extract the effect of gatekeeping by comparing the outcomes of people who are very similar in their observables, but enrolled in different insurance schemes. The second method is estimating the effect of gatekeeping on the initial provider choice of a sub-sample of respondents whose choice of health insurance plan was restricted by their employer offering only one health plan. As their enrollment into a particular type of insurance was driven by factors exogenous to their preferences over health care provision, no self-selection effect should be present.

3.2.1 Propensity score matching

Rosenbaum and Rubin (1983) proposed propensity score matching as a method to reduce selection bias in the evaluating of the treatment effects within the framework of non-randomized observational studies. This method is based on the idea that if one compares treatment and control groups that are as similar as possible in terms of their observable characteristics, it also reduces bias generated by unobservable factors. Therefore, the differences in outcomes for these two groups can be attributed to the effect of the treatment. In applying their idea to our selection

problem, the basic assumption would be that if the respondents are similar in their observables, they are also similar in their unobservable motives for the enrolling into the gatekeeping vs. non-gatekeeping plans. Therefore, conditional on the observables underlying the choice of insurance, the difference in the choice of ICP should be attributed only to the restrictions imposed by a given type of plan.

To implement this methodology, we used the preprogrammed STATA routine pscore[9] that estimates the propensity score according to the following algorithm.

1. Estimate a probit model of the choice of health insurance with gatekeeping restriction in the form

$$P(D_i = 1 \mid \boldsymbol{X}_i) = \Phi(h(\boldsymbol{X}_i)),$$

   where Φ denotes the normal cumulative distribution function, and $h(\boldsymbol{X}_i)$ is the starting specification of covariates. The basic specification of propensity score that we have chosen includes socio-demographic and employment characteristics as well as health indicators. Thus, it covers most of the insurance choice determinants pointed out in theory. To check the robustness of the results to the inclusion of different variables (as suggested in Caliendo and Kopeinig, 2008), we also estimated a second, more parsimonious specification, where we have selected only those covariates that are the best predictors of the gatekeeping status. Also, we restricted further analysis on the observations within the common support of the propensity score of gatekeeping and non-gatekeeping respondents.

2. Split the sample into an optimal number of intervals (blocks) by the propensity score such that within each block the mean propensity score does not differ between gatekeeping and non-gatekeeping individuals.

[9] For the detailed description of this routine, see Becker and Ichino (2002).

3. Within each interval, test whether the means of each characteristic do not differ between gatekeeping and non-gatekeeping individuals - i.e., test the balancing property. If the balancing property is violated, one has to look for another specification of propensity score.

As the probability of observing two individuals with exactly the same propensity score is, in principle, zero [since $p(\boldsymbol{X})$ is a continuous variable], one has to overcome this problem by designing a method to match people on the similarity of their propensity scores. We used three different matching estimators implemented in STATA - nearest neighbor, radius, and kernel matching[10] - to estimate the Average Treatment Effect on Treated (ATT) and then compare their results to assess the robustness of estimates. In principle, they differ in the method of choosing the control/pool of controls (non-gatekeeping) for the treatment (gatekeeping) observation. As stated in Caliendo and Kopeinig (2008), if the three methods give similar results, robustness is confirmed. If the results vary, on the other hand, we would need a further investigation to reveal the source of the disparity.

An additional problem arises when we want to compute standard errors of treatment effects and test for their statistical significance. The estimated variance should also include the variance due to the estimation of propensity score and account for restricting the estimation on common support. In our paper, we use the two most common alternatives in the applied literature – Lechner variance approximation (Lechner, 2001) and bootstrapping.

3.2.2 Estimation on the sub-sample with exogenous insurance choice

The second method of dealing with the endogeneity of gatekeeping status is to focus on the sub-sample who did not choose their insurance coverage and gatekeeping status themselves, and therefore, where this status should be unrelated to

[10] For a detailed description of the routines `attnd`, `attr`, and `attk`, see Becker and Ichino (2002).

their individual preferences or expectations about their future health care utilisation [see e.g . Martin, (1998)]. In our analysis, we focus on a sub-sample of covered respondents who were offered only one type of health insurance by their employer.

There are two implicit assumptions embedded in the application of this methodology. First, we assume that people do not choose their job based on the type of health insurance it offers. There are apparent differences in the characteristics of firms that offer none, single, or a broader choice of health plans. Naturally, the question arises as to whether different firms attract employees with varying health care preferences that will later translate into different patterns of utilisation (we try to answer this question in the Results section). Second, we assume that people generally prefer employer-provided insurance to the outside option of purchasing individual coverage, and therefore, the restricted choice is binding for them. This is true mainly due to financial concerns as in the US-employer provided insurance is much cheaper than individual coverage. Also in our data, we can see that almost 90 percent of the respondents who were offered some plan through their employer accepted it, and 95 percent of those who rejected are dependents on a family policy.

3.3 Further course of treatment

The episodic model of health care demand assumes that once the decision about the first point of contact has been made, patient delegates most of his decision-making authority over the further course of treatment to the initial contact provider. This provider is then responsible for directing the patient through treatment either by directly providing care or through referrals. Nevertheless, a patient's characteristics still influence health outcomes either because of their clinical importance or due to the patient's compliance with treatment.

Therefore, we develop our analysis of the ICP choice by focusing on how this choice affects the course of a patient's further treatment. The course of treatment is

estimated using two measures - the probability of further encounters with different types of providers within the episode of treatment (extensive margin) and the number of medical care events within the episode of treatment by type of provider (intensive margin). Control variables include socio-demographic characteristics, self-reported health status, and dummy variables for the most prevalent health conditions.

In our simple empirical model of the determinants of medical care utilization, the further course of treatment is assumed to depend on the actual choice of the initial contact provider. We also include the interaction of this choice with the gatekeeping restriction to estimate separately the utilization measures for people with and without gatekeeping restriction who initially choose the same type of provider. This way, we can identify whether any effect of gatekeeping existsthat stems from a source other than the initial provider choice.

In general, both measures of health care utilization are modeled through a density function f such as

$$P(Y_{ij} = y_{ij} \mid \boldsymbol{X}_i, d_i, p_{ik}) = f(\sum_k \alpha_k p_{ik} + \sum_k \delta_k p_{ik} d_i + \boldsymbol{X}_i' \boldsymbol{\beta}), \quad k = 1, \ldots, 5$$

where Y_{ij} denotes the utilization of services of provider type $j (j = 1, \ldots, 5)$[11] by individual i; $\boldsymbol{X}_i$ is the vector of independent explanatory variables; d_i is the binary indicator for the gatekeeping status; and p_{ik} are dummy variables indicating whether the provider of type k was the initial provider for individual i. The model of the probability of a further encounter is then specified as a simple logit model, while the number of visits to a particular type of provider, recorded as a non-negative integer count, is specified as a negative binomial-2 density to account for the excess number of zeros and over-dispersion.[12] We estimate the model on both

[11] Provider types are denoted as follows: 1 - primary care physician, 2 - specialist, 3 - non-physician medical personnel, 4 - hospital, and 5 - emergency room.

[12] We should note that the encounters with different types of medical care providers are not

the full sample and sub-sample with exogenous enrollment into health insurance plans so that we can assess the effect of self-selection into gatekeeping plans.

4 Data

4.1 General description

Data for our episode-level analysis are derived from the Household Component of the US Medical Expenditure Panel Survey (MEPS) for the years 2001 to 2006 (Panels 6, 7, 8, 9 and 10). These data are collected by the Agency for Health Care Research and Quality (AHRQ) and the National Center for Health Statistics by drawing a sample of households that participated in the previous year's National Health Interview Survey (NHIS) and then applying an overlapping panel design with five interviews occurring over a two-and-a-half year period. The sample is representative of the American civilian, non-institutionalized population with an oversampling of minorities[13].

MEPS data contain detailed information on medical expenditures and utilization, demographic characteristics, employment characteristics, health insurance coverage, and the health status of individuals. Moreover, MEPS groups medical care events (e.g. office visits, in-patient visits, or emergency room admissions) into episodes of care based on self-reported medical conditions, which enables us to use the treatment episode as the unit of analysis. Basically, we draw an individual and his characteristics from the MEPS Full Year Consolidated Data File, connect him with all of his reported medical conditions by a link to the MEPS Medical Conditions File, and get detailed information about all the medical care

mutually exclusive nor are they independent. A person with a severe condition would have a high probability of seeing more types of medical care providers in the course of his treatment. Therefore, one can consider estimating the equations for the utilization of different types of services as a system.

[13] For more information on MEPS survey design, see Cohen (1996, 1997) and Cohen (1997).

events connected to the conditions from the series of MEPS Event Files. This way, we avoid the main pitfall of studies using aggregate utilization measures over the whole period of the panel (i.e. undetected multiple illness spells) because we can connect treatment to a particular condition[14]. Also, we are able to detect the event that initiated the episode of care, which is crucial for our analysis of ICP choice.

In our analysis, we only consider medical care events identified as office visits, outpatient department visits, hospital inpatient stays, and emergency room visits[15]. Furthermore, we restrict our analysis to the first condition reported after June 1^{st} of the first year of the panel survey (e.g. year 2001 for respondents from Panel 6, year 2002 for respondents from Panel 7, etc.). The main reason is to avoid the problem of truncation. Since we do not observe any condition-related event within the first five-month period, we can assume that we have identified the true beginning of treatment and that the first reported event also represents the first contact for that episode. By including only one episode of treatment for each respondent, we ensure that all observations are independent. This approach also has drawbacks, however. We discard all people who do not report any medical care utilization, as well as those who have been treated only within the first five months of the survey. This strategy introduces a source of selection bias into our data, the magnitude of which will be discussed in the next sub-section. Also, we cannot confirm that the treatment episode had concluded by the end of the survey.

4.2 Sample construction

We focus on a sub-sample of non-elderly adults (ages 18-64) with private insurance only who have responded in all five interview rounds. The age restriction

[14] See Santos Silva and Windmeijer (2001) for a detailed discussion of the problem of multiple illness spells.

[15] We thus discard dental visits and home health care first because the dental coverage is usually separate from general coverage and second because home health care applies to long-term and chronic conditions that generally reach over the span of 2 years.

enables us to avoid selection bias originating in different age-specific insurance coverage opportunities: children are often covered by their parents' plans[16], and the elderly above the age of 64 are eligible for the publicly funded Medicare program. Also, these three groups have inherently different levels of medical utilization (unrelated to their specific insurance coverage) with children and the elderly tending to have higher medical care utilization than non-elderly adults. We further narrow the sample to employed (but not self-employed) individuals for whom we have information about employer characteristics. After dropping observations for which variables of interest were not defined, we are left with 18,809 observations.

We then merge this sample with information about the first condition reported after June 1^{st}, including the detailed characteristics of the ICP within this condition. This strategy resulted in dropping 6,020 respondents who did not report any medical care utilization (872 obs.) or all of their episodes of medical care utilization began before June 1^{st} (5,148 obs.). Thus, the final sample is 12,789 observations. Simple tabulation reveals that in the sample of dropped observations, we observe a significantly higher proportion of people with gatekeeping requirements. These are not the people who report no medical care utilization, however, as their distribution is the same across the two insurance types. Instead, respondents with gatekeeping restrictions tend to have fewer reported conditions in general and therefore a higher probability of reporting all medical care in the first five months of the survey.

4.3 Definition of variables

4.3.1 Insurance plans with gatekeeping restrictions

MEPS identifies the HMO and gatekeeper plans among privately covered individuals by asking a series of questions about the characteristics of the plan. First,

[16] Although young adults (students) are usually covered by their parents' insurance plans, they account only for 2.3 percent of the sample without choice of insurance. Also, the age division is standard in the literature.

the person is asked whether he is covered by an HMO. If the answer is negative, a follow-up question determines whether the person is in a gatekeeping plan other than an HMO. We use the answers provided in the first round to avoid the issue of the reverse causation, i.e., influence of changed health status on the choice of insurance. Consistent with prior studies (Pati, XXX, 2005), we then define gatekeeping enrollees as those who responded "yes" either to the first or the second question. From our tabulations, 59 percent of the sample has insurance with gatekeeping restrictions (hereafter referred to as "gatekeeping" enrollees) and 41 percent of the sample has insurance without gatekeeping restrictions (hereafter referred to as "non-gatekeeping" enrollees).

4.3.2 Medical care utilisation

The particular focus of this paper is on the first point of contact - i.e., the choice of ICP. We divide medical care providers into five categories – primary care physicians (PCP), specialists (SPEC), non-physician medical personnel (nonMD), hospitals (HOSP), and emergency rooms (ER). A physician was designated as a PCP if he or she was a general or family practitioner or general internist[17]. Any other medical field was designated as "specialist".

Table 4.1 presents a summary of utilization measures by gatekeeping status. First, we present general summary measures of the number of conditions reported as well as total numbers of visits to a particular provider, all as a total over the two years of the panel survey. In general, gatekeeping enrollees report fewer conditions and fewer non-MD encounters. With respect to other types of medical care utilization, we do not see significant difference between gatekeeping and non-gatekeeping enrollees. We also report what we refer to as health indicators, i.e. responses concerning the individual's health status in the first round of interviews, which could

[17]Note that we did not include obstetricians/gynecologists in the category of PCP's but as specialists instead . Even so, this definition tends to overestimate PCPs in the non-gatekeeping plan and therefore any differences in PCP use between the two plans would be underestimated.

predict further medical care utilization. We include dummies for being healthy [18], for physical limitations and smoking, and Body Mass Index (BMI). The comparison shows that a significantly lower percentage of gatekeeping respondents consider themselves to be healthy; other health indicators, however, do not appear to be correlated with the gatekeeping status.

The second part of Table 4.1 describes utilization within the first observed condition after June 1^{st}, in terms of the ICP chosen as well as a summary of further utilization measures. A summary of the first encounter gives us the first insight into the question: "How does the gatekeeping requirement affect the choice of the initial contact provider?" Simple tabulation suggests that gatekeeping enrollees have a significantly higher probability of visiting a PCP and a lower probability of visiting a specialist as their first point of contact than non-gatekeeping enrollees. This difference is not as striking as one would expect, however, under such an explicit restriction: only 3 percentage points in the case of PCP and 3 percentage points in the case of specialist visits. Also, in terms of further utilization, gatekeeping and non-gatekeeping enrollees are very similar with the exception that gatekeeping respondents have on average more PCP visits in the course of their treatment.

Finally, we control for the particular conditions reported, which were chosen either due to their prevalence in the sample or due to the specific and predetermined course of treatment in terms of provider choice. These conditions are hypertension, upper respiratory infection, pregnancy, intervertebral disc dislocation, sprains and strains, wounds, other injuries, joint disorders, connective tissue disease, skin disorders, diabetes, neoplasm, lipid metabolism disorders, blindness, chronic pulmonary conditions, intestinal infections, and urinary calculus. Their prevalence among gatekeeping and non-gatekeeping respondents is summarized in Table 4.2.

[18] This dummy was derived from the self-reported perceived health status variable ascertained in the first interview round, where we designated respondents who answered "excellent" and "very good" as healthy.

4.3.3 Other covariates

Other covariates used in the estimation are summarized in Table 4.3. We divide them into two categories: socio-demographic and employment-related variables.

Socio-demographic characteristics include age, sex, race, years of education by the time of entering MEPS (top coded at 17), region of residence, urban status (whether the person resides in a metropolitan statistical area), marital status, family size (number of children), and the natural logarithm of family income. When compared to non-gatekeeping enrollees, gatekeeping enrollees have a higher probability of being a minority (hispanic, black, or Asian), fewer years of schooling, have a higher probability of living in a city, are less likely to be married, and have bigger families with more children.

Employment characteristics and the availability of health insurance through an employer are interesting variables because they are assumed to be determinants of the choice of health plans, but should not affect gatekeeping enrollment. Employment characteristics include number of employees at the current job location, an indicator for being employed by a small company (1-10 employees), an indicator for firms with more locations, and union status. Furthermore, the MEPS includes information on whether or not the employer offers a health insurance plan; whether, conditional on this offer, he provides a choice of plans or just one alternative; and whether the employee eventually holds insurance through his employer. Indeed, all these characteristics vary significantly by gatekeeping status, with gatekeeping enrollees having a higher probability of working in larger companies with a choice of health plans than non-gatekeeping enrollees.

5 Results

5.1 Choice of the initial provider

5.1.1 Propensity score matching

Table 5.1 presents the estimation results of the propensity score of being enrolled in the gatekeeping program. We estimated two propensity score specifications, the first using all relevant socio-demographic and employment related variables and the second using only variables identified as best predictors of gatekeeping status. We also present a graphic illustration of the distribution of gatekeeping and non-gatekeeping enrollees by their estimated propensity score (Figures 5.1 and 5.2). Since the common support encompasses almost the whole population of gatekeeping and non-gatekeeping respondents, we have sufficient overlap to estimate treatment effects.[19]

Comparison of the estimates of the gatekeeping's effect on the choice of initial contact provider is presented in Table 5.2. We compare the results of propensity score matching to the estimates from the comparison of unmatched samples, where the assignment into gatekeeping status is assumed to be exogenous. For each type of ICP and corresponding matching estimator, we report the number of treated and control respondents used, estimated Average Treatment Effect on Treated (ATT), standard errors and t-statistics derived from Lechman variance formula, and standard errors and 99 percent confidence intervals from bootstrapping.

In general, propensity score matching estimates show a positive statistically significant effect of gatekeeping on the probability of choosing a PCP as the ICP (approx. 3.7 percentage points) and a negative significant effect on the probability of choosing a specialist as the ICP (approx. 4 percentage points). These methods

[19] One should notice, however, that although both specifications passed the likelihood ratio test of the coefficients' significance, we are not able to predict gatekeeping status with a high accuracy (we obtained pseudo R-squared of 0.05).

find no significant effect of gatekeeping on the probability of choosing non-MD personnel, hospital, or emergency room as the ICP. The results are similar to the estimates on the unmatched sample both in terms of direction and magnitude, which would suggest little selection on observables.

As a robustness check, we compare the results from two different specifications and three different matching mechanisms. In comparing the two specifications, the results are generally consistent, but the second specification provides slightly higher estimates of the gatekeeping effects. In comparing different matching methods, results are fairly robust, with the exception of the nearest neighbor estimator used in the first specification, which consistently gives us slightly higher estimates than the other two methods. In this particular case, we believe it results fromthe lower number of controls used. This method uses only two-thirds of the non-gatekeeping sub-sample. We thus consider the other two methods more reliable.

5.1.2 Estimation on sub-sample with exogenous insurance choice

We implement the methodology outlined in section 3.2.2 by restricting the sample to respondents who held a health insurance policy provided by their employer that was the only option offered by an employer. We excluded respondents who at any time during the reference period acquired an additional health insurance policy, either as a policy holder (e.g. to cover special health care requirements) or as a dependent (e.g. within family coverage).

First, we check whether the sub-sample is systematically different from the full sample in terms of individual characteristics and health care utilization. In Table 4.3, panel (2), we summarize the socio-demographic and employment-related characteristics. Our sub-sample consists of people who are slightly older and less educated than the average respondent, with a higher proportion of males. The ethnic and regional distributions, as well as type of marital status and family income

are similar to the full sample. Respondents from this sub-sample typically work in firms with fewer employees but not in the smallest firms (with fewer than 10 employees). This finding is consistent with the observation that medium-sized firms usually offer health insurance coverage but with a limited choice of plans.

The main difference in health insurance between the full sample and sub-sample is that the sub-sample has a much lower share of gatekeeping enrollees (53 percent in the sub-sample compared to 61 percent in the full sample). This finding suggests that when people can choose their coverage type, they opt for a gatekeeping plan. Moreover, as shown in Table 4.1, panel (2), respondents in the sub-sample have lower levels of general medical care utilization, with the exemption of primary care utilization. Summary characteristics of the first event within the selected medical condition shows that the ICP choice, as well as pattern of further utilization, is similar between the full sample and sub-sample. While in the full sample the main difference between the gatekeeping and non-gatekeeping enrollees is in the choice between PCPs and specialists in the sub-sample, the substitution happens between PCPs and non-medical personnel.

Table 5.3 summarizes results from estimating a linear probability model (standard OLS) for the choice of ICP within the first medical condition that occurred after 5 months of monitoring for both full sample and sub-sample. We include two broad classes of exogenous covariates that influence the choice of the initial provider – condition indicators and demographic characteristics (sex, race/ethnicity, education, region, marital status, income, number of children, and self-preceived health status). The OLS results for the full sample suggest that gatekeeping restrictions have a positive and statistically significant effect on the probability of contacting a PCP (3.6 percentage points) and a negative, significant effect on the probability of contacting a specialist (3 percentage points). On the other hand, estimates performed on the sub-sample indicate only a weak, positive effect of gatekeeping

on the probability of first contact with a PCP (3.2 percentage points at 10 percent significance level), while having no statistically significant effect on the probability of contacting a specialist or any other type of providers. To put these results into perspective: the share of patients who self-refer themselves decreased from 24 percent to 21 percent for the full sample and remains at 22-23 percent for the sub-sample. The influence of the gatekeeping restriction on the ICP choice is therefore small.

As for the estimated coefficients on the other explanatory variables, condition indicators are, in general, very good predictors of initial provider choice. Also, they are similar both in sign and magnitude over the full sample and sub-sample. This finding is probably due to our choice of conditions with fairly standardized courses of treatment. Demographic characteristics, in general, do not seem to affect the probability of hospitalization and have a relatively low effect on the probability of a PCP visit, but they are a significant predictor of choosing a specialist, non-medical personnel, or ER as the ICP. This pattern also holds for the sub-sample, with the exemption of encounters with specialists, where the demographic variables lose their explanatory power and choice is primarily determined by the type of condition.

5.2 Course of further treatment

In this section, we explore further possible mechanisms of gatekeeping's indirect impact by analyzing the effect of IPC choice on patterns of further medical care utilization. The results of the simple model of further utilization, estimated on the full sample, are presented in Tables 5.4 - 5.8. Tables 5.4 and 5.5 summarize the average sample probabilities and numbers of encounters with a given type of provider during the episode of care conditional on the choice of ICP and gatekeeping status. Table 5.6 then contains the results of estimation of a logit model for the probabilities of an encounter, while Table 5.7 contains results of estimation of a negative

binomial regression model for the number of encounters. In both tables, columns represent utilization outcomes of interest (probability and number of encounters with a given type of provider), while rows represent the marginal effects of binary indicators of initial provider status in interaction with the gatekeeping status of the respondent with a PCP contacted by a non-gatekeeping enrollee being the base category. Both types of models were estimated with and without additional covariates, corresponding to the first and second column within each provider category. We do not report the outcomes for other included covariates, but we comment on their significance in particular cases later.

The results of the estimation lead us to two basic conclusions: (1) we can observe distinctly different patterns of medical care utilization conditional on the choice of ICP; and (2) after controlling for ICP choice, gatekeeping requirements in general do not have any additional impact on further utilization, with a few exceptions. These results hold for both extensive and intensive utilization measures, are statistically significant, and are robust to the inclusion of other covariates.

More detailed analysis of utilization measures reveals other interesting patterns. PCP-initiated episodes of care have the lowest further utilization measures both in terms of probability and number of events. On the other hand, episodes initiated by a visit to a specialist have a high probability of continuing treatment by the specialist and also have an increased probability of ending in a hospital, which means that specialists see enrollees with more serious conditions.

From the theory, we can infer that two types of patients are induced to switch from secondary to primare care by the gatekeeping restriction. First, there are patients who do not need specialist care, and PCPs contain costs by keeping them within primary care. Second, there are patients who need to see a specialist for whom the initial visit to a PCP results in an immediate referral. In the comparison of gatekeeping and non-gatekeeping patients, we observe that gatekeeping patients

have a significantly higher probability of referral to a specialist[20]. Thus, we can infer that the second type of patient prevails. [21] At this point, we cannot make any inference regarding the effectiveness of gatekeeping in such a setting, however, since these patients could be either those who would self-refer correctly and for whom the initial PCP visit was redundant, or patients who did not know the proper specialist and benefited from the screening provided by a PCP.

Episodes initiated by a visit to non-MD personnel have a very specific character. They lead to a high probability of continuing treatment by non-MD personnel with multiple visits (8 times more visits as compared to when the treatment was initiated by a PCP visit). From the data, we indicate that these are the episodes connected to specific chronic conditions with a standardized treatment procedure.[22]

If the episode starts with a hospital admission, the expected future medical care utilization is on average highest among the alternatives, which suggests that hospitalization can be used as a proxy for illness severity. Finally, emergency room visits often result in further treatment by specialists or even hospitalisation, but there is no significant difference between the outcomes of gatekeeping and non-gatekeeping enrollees. Thus, we confirm that ERs do not provide after-hours care for patients with gatekeeping insurance plans, as is the case with uninsured patients.

5.2.1 Appropriateness of self-referral

Self-referral is a common path to specialist care. Based on our summary tabulations, 24 percent of non-gatekeeping and 21 percent of gatekeeping enrollees self-refer. The effectiveness and appropriateness of self-referral depends on the patient's

[20]Their probability of being referred to a specialist is 2.5 percentage points higher, which represents a 24.5 percent increase.

[21]This observation was confirmed by a repeated estimation on the sub-sample of respondents with exogenous choice of health insurance for which we did not find significant effect of gatekeeping on ICP choice in the first stage. Consistently with our inference, we also did not find any difference in PCP-referral rates between gatekeeper and non-gatekeeper patients.

[22]Types of medical personnel contacted the most are chiropractors, nurses, technicians, physical therapists and psychologists. Conditions that are treated by non-medical personnel are intervertebral disc dislocation, connective tissue disease, and sprains and strains.

ability to assess the severity of his condition and choose an appropriate provider type. Inspired by Forrest and Reid (1997), we try to evaluate the appropriateness by using a simple proxy for condition severity, i.e. hospitalization.

We re-estimate the logit model for the probability of hospitalization on the subsample of respondents who either have visited a PCP as the ICP and were then referred to a specialist or have self-referred to specialist directly. First, we examined whether the probability of hospitalization varies by type of ICP. Our findings are consistent with the previous literature. In the simplest version of the model, i.e. estimation after controlling for individual-specific covariates (not presented in the table), respondents who self-referred to specialty care had a 4 percentage point lower probability of being hospitalized than respondents who were referred by a PCP. With a baseline hospitalization rate of 0.073 for PCP-referred patients, this corresponds to a 55 percent lower hospitalization rate.

Further, we add interaction terms in order to estimate separate effects for gatekeeping and non-gatekeeping enrollees. Table 6.8 summarizes the results of estimation both on the full sample of PCP-referred and self-referred patients, as well as exclusively on those who were defined as having no choice of insurance plan, according to section 5.1.2. We see that the difference in hospitalization rates is based on the difference within gatekeeping enrollees. While there is no significant difference in the hospitalization rates within non-gatekeeping enrollees, there is an almost 6 percentage point difference between the hospitalization rate of PCP-referred and self-referred within gatekeeping enrollees. For the sample of respondents without insurance choice the difference increases to 9.4 percentage points, i.e. the chances of hospitalization for PCP-referred patients are 3.3 times greater than for the self-referred!

6 Conclusion

The majority of current studies on the effectiveness of managed care evaluate the role of gatekeeping restrictions by estimating their effect on aggregate measures of health care utilization. This approach is questionable, however, because gatekeeping is primarily intended as a mechanism to reinforce the use of primary care physicians as initial contact providers. Taking advantage of individual-level panel data on medical care utilization from 2001 - 2006 Medical Expenditure Panel Survey, we are able to isolate multiple effects of gatekeeping restrictions including their influence on the choice of initial care providers and on the course of further treatment. We explore a quasi-natural experiment in our data arising from differences in the degree of choice of insurance plans employers offer their employees to account for the effect of self-selection into gatekeeping plans.

Our main results are counterintuitive and genuinely surprising. The theory behind the concept of gatekeeping predicts fewer self-referrals to specialists and a corresponding increase in the number of primary care physician visits for individuals with gatekeeping requirements as opposed to those enrolled in non-gatekeeper plans. Nevertheless, our results show only economically trivial (although statistically significant) differences. Probing more deeply, we see that most of the patients who were induced to use their primary care physicians as initial contact providers are referred back to a specialist, i.e. they indeed needed specialty care. This finding, together with the fact that within gatekeeping plans 21 percent of patients still self-refer to specialists, implies that the intended economic effect of gatekeeping, reducing utilization of specialty care, is surprisingly weak.

When we assess the appropriateness of self-referrals, we find that self-referred patients are less severely ill than patients who were referred to specialists by primary care physicians. Two mechanisms can explain these results. The first is the behavior of patients. While without gatekeeping restriction patients with severe illnesses

seem to distribute themselves randomly between PCP and specialist, they seem to behave differently under the gatekeeping restriction. They self-refer mainly with regular and recognizable conditions that need specialist attention but are only moderately severe, while with any other condition they visit their PCP. The second mechanism is the screening role of PCPs. As the first-contact provider, they are retaining less sick patients in a primary care setting while referring the more severely ill to specialty care. This tendency towards screening is generally incentivized in gatekeeping plans (by, for example, provisions that limit physician referral rates).

Therefore, it appears that gatekeeping operates through channels other than those typically assumed. It does not affect direct access to specialty care as much as it changes the composition of patients who self-refer and patients who are referred by a primary care physician. This has important implications for the designers of health insurance policies as it implies differentiated effect of ga atekeeping restriction on the agents within a health system. On one hand, the behavior of patients seems to be only slightly modified by gatekeeping. On the other hand, the incentives for gatekeeping primary care physicians can induce a higher efficiency of screening and treatment process.

References

[1] Arrow, K.J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review* 53, 941-973.

[2] Atella, V. and Deb, P. (2008). Are Primary Care Physicians, Public and Private Sector Specialists Substitutes or Complements? Evidence from Simultaneous Equations Model for Count Data. *Journal of Health Economics* 27 (3), 770-785.

[3] Becker, S.O. and Ichino, A. (2002). Estimation of Average Treatment Effects Based on Propensity Scores. *The Stata Journal* 2(4), 358 - 377.

[4] Caliendo, M. and Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys* 22 (1), 31-72.

[5] Campbell et al. (2007). Professionalism in Medicine: Results of a National Survey of Physicians. Annals of Internal Medicine 147, 795-802.

[6] Cohen, S. B. (1996). The Redesign of the Medical Expenditure Panel Survey: A Component of the DHHS Survey Integration Plan. Proceedings of the COPAFS Seminar on Statistical Methodology in the Public Service.

[7] Cohen, S. B. (1997). A Sample Design of the 1996 Medical Expenditure Panel Survey Household Component, Rockville (MD): Agency for Healthcare Research and Quality; 1997. MEPS Methodology Report, No. 2. AHCPR Pub. No. 97-0027.

[8] Cohen, J. W. (1997). A Design and Methods of the Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Healthcare Research and Quality; 1997. MEPS Methodology Report, No.1. AHCPR Pub. No. 97-0026.

[9] Deb, P. and Trivedi, P.K. (2006a). Provider Networks and Primary Care Signups: Do They Restrict the Use of Medical Services? Hunter College CUNY, Working paper 2004.

[10] Deb, P. and Trivedi, P.K. (2006b). Restrictions on Provider Access in Health Plans and Socioeconomic Status. *Health Services Research* 41(5), 1821 - 1846.

[11] Donaldson, M.D. et al. (1996). Primary Care: America's Health in New Era. in: Donaldson, M.S. et al. (Eds.) Institute of Medicine, Division of Health Care Services, Committee on the Future of Primary Care. National Academy Press, Washington D.C.

[12] Forrest, Ch.B., Nutting, P.A., Starfield, B. and von Schrader, S. (2002). Family Physicians' Referral Decisions. *Journal of Family Practice* 51(3), 215 - 222.

[13] Forrest, Ch. B. and Reid, R.J. (1997). Passing The Batton: HMO's Influence on Referrals to Specialty Care. *Health Affairs* 16 (6), 157 -162.

[14] Fortney, J.C, Steffick, D.E., Burgess Jr., J.F., Maciejewski, M.L. and Petersen, L.A. (2005). Are Primary Care Services a Substitute or Complement? *Health Services Research* 40:5, Part I, 1422-1442.

[15] Glied, S. (2000). Managed care. in: Culyer, A.J. and Newhouse, J.P. (Eds.) Handbook of Health Economics, Vol 1A . North - Holland, Amsterdam, 707-745.

[16] Grembowski, D.E., Cook, K. and Patrick, D.L. (1998) Managed care and physician referral. *Medical Care Research and Review* 55 (1), 3-31.

[17] Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *The Journal of Political Economy* 80(2), 223-255.

[18] Hellinger, F.J. (1995). Selection bias in HMOs and PPOs: A Review of the evidence. *Inquiry* 32 (4), 135-42.

[19] Holmes, A.M. and Deb, P. (1998). Provider Choice and Use of Mental Health Care: Implications for Gatekeeper Models. *Health Services Research* 33(5), 1263 - 1284.

[20] Kassirer, J.P. (1994). Access to Specialty Care. *New England Journal of Medicine* 331(17), 1151-1153.

[21] Lechner, M.(2001). Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption. in: Economic Evaluation of Labour Market Policies, eds. Lechner, M. and F. Pfeiffer, Heidelberg: Physica, 1-18.

[22] Martin, D.P., Diehr, P., Price, K.F. and Richardson, W.C. (1989). Effect of a gatekeeper plan on health services use and charges: a randomized trial. *American Journal of Public Health* 79 (12), 1628- 1632.

[23] Miller, R.H. and Luft, H.S. (1997). Does managed care lead to better or worse quality care? *Health Affairs* 16 (5), 7-25.

[24] Miller, R.H. and Luft, H.S. (2002). HMO plan performance update: an analysis of the literature, 1997 - 2001. *Health Affairs* 21(4), 63-86.

[25] Mullahy, J. (1997). Instrumental Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior. *The Review of Economcs and Statistics*, 79(4), 586-593.

[26] Pati, S., Shea, S., Rabinowitz, D. and Carrasquillo, O. (2005). Health Expenditures for Privately Insured Adults Enrolled in Managed Care Gatekeeping vs. Indemnity Plans. *American Journal of Public Health*, 95:2, 286-291.

[27] Pauly, M.V. (1968). The Economics of Moral Hazard: Comment. *American Economic Review* 58, 531-537.

[28] Pohlmeier, W. and Ulrich, V. (1995) An Econometric Model of Two-part Decision Making Process in the Demand for Health Care. *Journal of Human Resources* 30 (2): 339-361.

[29] Rosenbaum, P.R. and Rubin, D.B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70, 41-55.

[30] Santos Silva, J.M.C. and Windmeijer, F. (2001). Two-part multiple spell models for health care demand. *Journal of Econometrics* 104, 67-89.

[31] Scott, A. (2000). Economics of general practice, Chapter 22 in: Culyer, A.J. and Newhouse, J.P. (Eds.) Handbook of Health Economics, Vol. 1A North - Holland, Amsterdam, 1175 -1200.

[32] Starfield, B. (1994). Is primary care essential? *The Lancet* 344, 1129-1133.

[33] Stoddart, G. and Barer, M. (1981). Analyses of Demand and Utilization Through Episodes of Medical Service. in: J. van der Gaag and Perlman, M. (Eds.) Health, Economics, and Health Economics. New York: North Holland, 149-70.

[34] Zweifel, P.(1981). Supplier-Induced Demand in a Model of Physician Behavior. in: Health, Economics and Health Economics, ed. Jacques van der Gaag and Morris Perlman., Amsterdam: North-Holland, 245-67.

Table 4.1: Summary of medical care utilisation measures for (1) the full sample and (2) the sub-sample of respondents with exogenous choice of health insurance plan by gatekeeping status

	(1) Full sample				(2) Sub-sample			
	non-gatekeeping		gatekeeping		non-gatekeeping		gatekeeping	
	mean	SE	mean	SE	mean	SE	mean	SE
General utilization								
# of conditions	3.93**	[0.04]	3.81**	[0.03]	3.79	[0.07]	3.69	[0.07]
total # of PCP visits	3.92	[0.10]	3.92	[0.07]	4.10	[0.24]	4.12	[0.16]
total # of SPEC visits	4.83	[0.11]	4.82	[0.11]	4.40	[0.21]	4.27	[0.22]
total # of nonMD visits	7.32***	[0.25]	6.24***	[0.18]	6.38	[0.40]	5.82	[0.41]
total # of HOSP visits	0.24	[0.01]	0.24	[0.01]	0.19	[0.01]	0.21	[0.02]
total # of ER visits	0.45	[0.01]	0.44	[0.01]	0.39	[0.02]	0.44	[0.03]
Health indicators								
healthy (0/1) [b)]	0.66***	[0.01]	0.62***	[0.01]	0.63***	[0.01]	0.58***	[0.01]
limitations (0/1) [a)]	0.07	[0.01]	0.08	[0.01]	0.07	[0.01]	0.08	[0.01]
BMI [a)]	27.9	[0.08]	27.8	[0.07]	28.6	[0.16]	28.2	[0.15]
smoking [a)]	0.20	[0.01]	0.19	[0.01]	0.21	[0.01]	0.22	[0.01]
First condition after June 1^{st}								
First encounter								
PCP	0.43***	[0.01]	0.46***	[0.01]	0.44**	[0.01]	0.48**	[0.01]
specialist	0.24***	[0.01]	0.21***	[0.01]	0.23	[0.01]	0.22	[0.01]
nonMD	0.20	[0.01]	0.19	[0.01]	0.20**	[0.01]	0.17**	[0.01]
hospital	0.02	[0.01]	0.02	[0.01]	0.02	[0.01]	0.02	[0.01]
emergency	0.11	[0.01]	0.11	[0.01]	0.11	[0.01]	0.12	[0.01]
Further utilization								
# of PCP visits	0.82**	[0.02]	0.88**	[0.02]	0.97	[0.07]	0.93	[0.06]
# of SPEC visits	1.00	[0.03]	1.00	[0.03]	0.81*	[0.04]	0.91*	[0.04]
# of nonMD visits	1.28	[0.07]	1.15	[0.05]	1.3**	[0.14]	0.96**	[0.08]
# of HOSP visits	0.07	[0.01]	0.06	[0.01]	0.06	[0.01]	0.05	[0.01]
# of ER visits	0.15	[0.01]	0.15	[0.01]	0.15	[0.01]	0.14	[0.01]
N	5203		7586		1397		1576	

Note: The differences between gate and non-gatekeeping enrollees significant at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.
[a)] The averages made over the sub-sample of respondents that were eligible and responded to the Self-Administered Questionnaire - approx. 87% of the sample.
[b)] The averages made over the sub-sample of respondents who answered the question.

Table 4.2: Prevalence of selected health conditions by gatekeeping status

	non-gatekeeping		gatekeeping		
Conditions	mean	SE	mean	SE	t-stat
hypertension	0.03	[0.003]	0.04	[0.002]	−0.57
upper resp. infection	0.07	[0.004]	0.07	[0.003]	1.30
pregnancy[a)]	0.04	[0.004]	0.05	[0.003]	−0.29
disc dislocation	0.05*	[0.003]	0.04*	[0.002]	1.87
sprains & strains	0.04	[0.003]	0.03	[0.002]	0.64
joint disorders	0.03	[0.002]	0.03	[0.002]	−0.66
connective tissue	0.03	[0.002]	0.04	[0.002]	−1.68
skin disorders	0.05	[0.003]	0.05	[0.002]	0.05
other injuries	0.02	[0.002]	0.03	[0.002]	−0.94
diabetes	0.01*	[0.001]	0.01*	[0.001]	−0.66
neoplasm	0.02***	[0.002]	0.01***	[0.001]	2.60
lipid metabolism	0.02	[0.002]	0.02	[0.002]	0.03
blindness	0.01	[0.002]	0.02	[0.001]	−1.06
chronic pulmonary cond.	0.02	[0.002]	0.02	[0.002]	0.19
intestinal infection	0.02	[0.002]	0.02	[0.002]	−1.70
urinary calculus	0.01	[0.001]	0.01	[0.001]	1.00
wounds	0.01	[0.002]	0.01	[0.001]	0.49

Note: The differences between gate and non-gatekeeping enrollees significant at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.
[a)] The prevalence calculated over the sub-sample of women.

Table 4.3: Socio-demographic and employment-related characteristics of (1) the full sample and (2) the sub-sample of respondents with exogenous choice of health insurance plan by gatekeeping status

	(1) Full sample				(2) Sub-sample			
	non-gatekeeping		gatekeeping		non-gatekeeping		gatekeeping	
	mean	SE	mean	SE	mean	SE	mean	SE
Demography								
age	41.8	[0.162]	41.9	[0.106]	43.5	[0.286]	43.7	[0.269]
male	0.44	[0.006]	0.43	[0.006]	0.51	[0.013]	0.50	[0.013]
years of education	13.68**	[0.034]	13.56**	[0.031]	13.32***	[0.067]	13.03***	[0.071]
- ***race/ethnicity dummies***								
hispanic	0.09***	[0.004]	0.15***	[0.004]	0.10***	[0.008]	0.19***	[0.010]
black	0.11***	[0.004]	0.12***	[0.003]	0.11	[0.008]	0.12	[0.008]
white	0.77***	[0.006]	0.68***	[0.005]	0.77***	[0.011]	0.66***	[0.012]
asian	0.03***	[0.002]	0.04***	[0.002]	0.02*	[0.004]	0.03*	[0.005]
-***region dummies***								
northeast	0.12***	[0.005]	0.20***	[0.005]	0.11***	[0.008]	0.20***	[0.010]
midwest	0.31***	[0.006]	0.22***	[0.005]	0.31***	[0.012]	0.21***	[0.010]
south	0.40***	[0.007]	0.33***	[0.005]	0.45***	[0.013]	0.37***	[0.012]
west	0.16***	[0.005]	0.25***	[0.005]	0.14***	[0.009]	0.21***	[0.010]
urban status (0/1)	0.73***	[0.006]	0.85***	[0.004]	0.68***	[0.012]	0.82***	[0.010]
-***marital status dummies***								
married	0.67***	[0.006]	0.64***	[0.006]	0.65***	[0.013]	0.59***	[0.012]
divorced	0.12**	[0.004]	0.13	[0.004]	0.16	[0.010]	0.17	[0.010]
single	0.18	[0.005]	0.19**	[0.005]	0.16*	[0.010]	0.17	[0.010]
family size	2.87***	[0.019]	2.95***	[0.017]	2.82	[0.037]	2.83	[0.038]
log(income)	10.9	[0.010]	10.9	[0.007]	10.90*	[0.017]	10.86*	[0.015]
# of children	0.81**	[0.015]	0.85**	[0.013]	0.80	[0.019]	0.79	[0.028]
Employment								
# of employees	168.9***	[2.6]	186.2***	[2.2]	157.1	[4.7]	165.1	[4.6]
small firm (0/1)	0.17***	[0.005]	0.14***	[0.004]	0.15*	[0.010]	0.13*	[0.009]
more locations[a)] (0/1)	0.71**	[0.006]	0.73**	[0.005]	0.67	[0.013]	0.67	[0.013]
unionized[a)] (0/1)	0.16	[0.005]	0.17	[0.004]	0.20	[0.011]	0.17	[0.009]
offer insurance (0/1)	0.85***	[0.005]	0.88***	[0.004]	–	–	–	–
choice of plans (0/1)	0.49***	[0.008]	0.58***	[0.006]	–	–	–	–
hold insurance (0/1)	0.75***	[0.006]	0.78***	[0.005]	–	–	–	–

Note: The differences between gate and non-gatekeeping enrollees significant at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.
[a)] The averages made over the sub-sample that responded who answered the question.

Table 5.1: Propensity score for enrollment into gatekeeping health insurance plan using probit

Specification 1

	Coef.	SE	z-stat
age	.0004	[0.001]	0.33
male	−.031	[0.023]	-1.31
hispanic	.146	[0.069]	2.11
black	.006	[0.069]	0.09
white	−.163	[0.061]	-2.67
years of education	−.007	[0.005]	-1.44
northeast	.476	[0.005]	13.12
west	.416	[0.034]	12.05
south	.041	[0.029]	1.39
married	−.089	[0.033]	-2.71
divorced	.035	[0.043]	0.81
family size	.047	[0.015]	3.08
# of children	−.002	[0.019]	-1.61
healthy	−.070	[0.024]	-2.88
choice of plans	.204	[0.025]	8.24
hold insurance	.012	[0.029]	0.42
# of employees	.0001	[0.001]	0.92
small firm	−.065	[0.034]	-1.91
constant	.131	[0.112]	1.17
Log Likelihood	=	−8329	
Numbers of obs	=	12789	
LR chi2(22)	=	623.66	
Pseudo R2	=	0.036	

Specification 2

	Coef.	S.E.	z-stat
urban status	.370	[0.035]	12.80
choice of plans	.167	[0.023]	7.14
small firm	−.072	[0.032]	-2.28
midwest	−.417	[0.029]	-14.14
south	−.351	[0.027]	-13.07
hispanic	.268	[0.036]	7.40
constant	.081	[0.033]	2.42
Log Likelihood	=	−8286	
Numbers of obs	=	12789	
LR chi2(22)	=	711.64	
Pseudo R2	=	0.041	

Table 5.2: Estimation of the effect of gatekeeping restriction on the probability of the choice of the initial contact provider using propensity score matching

Specification 1

	# treat	# contr	ATT	SE_Lech	t_Lech	SE_boot	99% CI	
PCP	unmatched		**0.037*****	[0.009]	3.74			
nearest neighbor	7586	3264	**0.042*****	[0.012]	3.43	[0.015]	0.022	0.068
radius matching	7583	5202	**0.037*****	[0.009]	4.37	[0.009]	0.006	0.055
kernel matching	7596	5202	**0.037*****	.	.	[0.014]	0.003	0.068
SPEC	unmatched		**−0.031*****	[0.008]	−3.92			
nearest neighbor	7586	3264	**−0.041*****	[0.011]	−1.65	[0.011]	−0.055	−0.021
radius matching	7583	5202	**−0.035*****	[0.008]	−4.48	[0.011]	−0.058	−0.003
kernel matching	7596	5202	**−0.033*****	.	.	[0.011]	−0.052	−0.007
NONMD	unmatched		**−0.007**	[0.007]	−0.37			
nearest neighbor	7586	3264	**−0.001**	[0.010]	−0.42	[0.011]	−0.029	0.024
radius matching	7583	5202	**−0.006**	[0.007]	−0.77	[0.009]	−0.026	0.021
kernel matching	7596	5202	**−0.002**	.	.	[0.010]	−0.025	0.023
HOSP	unmatched		**−0.002**	[0.002]	−1.32			
nearest neighbor	7586	3264	**−0.005**	[0.003]	−1.366	[0.003]	−0.019	−0.005
radius matching	7583	5202	**−0.002**	[0.002]	−1.01	[0.004]	−0.010	0.001
kernel matching	7596	5202	**−0.002**	.	.	[0.004]	−0.015	0.001
ER	unmatched		**0.003**	[0.006]	0.39			
nearest neighbor	7586	3264	**0.005**	[0.008]	0.572	[0.008]	−0.011	0.023
radius matching	7583	5202	**0.003**	[0.006]	0.51	[0.008]	−0.02	0.013
kernel matching	7596	5202	**−0.000**	.	.	[0.008]	−0.022	0.015

Specification 2

PCP	unmatched		**0.037*****	[0.009]	4.12			
nearest neighbor	7586	5202	**0.036*****	[0.010]	3.69	[0.010]	0.017	0.053
radius matching	7586	5202	**0.041*****	[0.010]	4.22	[0.009]	0.023	0.067
kernel matching	7586	5202	**0.037*****	.	.	[0.011]	0.020	0.056
SPEC	unmatched		**−0.031*****	[0.008]	−4.09			
nearest neighbor	7586	5202	**−0.039*****	[0.008]	−4.72	[0.008]	−0.054	−0.025
radius matching	7586	5202	**−0.045*****	[0.008]	−5.39	[0.010]	−0.062	−0.029
kernel matching	7586	5202	**−0.039*****	.	.	[0.010]	−0.065	−0.024
NONMD	unmatched		**−0.007**	[0.007]	−0.93			
nearest neighbor	7586	5202	**−0.002**	[0.008]	−0.25	[0.008]	−0.019	0.010
radius matching	7586	5202	**−0.004**	[0.008]	−0.56	[0.008]	−0.019	0.011
kernel matching	7586	5202	**−0.003**	.	.	[0.008]	−0.022	0.011
HOSP	unmatched		**−0.002**	[0.002]	−1.02			
nearest neighbor	7586	5202	**−0.001**	[0.003]	−0.49	[0.002]	−0.006	0.002
radius matching	7586	5202	**−0.001**	[0.003]	−0.45	[0.002]	−0.006	0.002
kernel matching	7586	5202	**−0.001**	.	.	[0.002]	−0.007	0.003
ER	unmatched		**0.003**	[0.006]	0.51			
nearest neighbor	7586	5202	**0.006**	[0.006]	1.07	[0.006]	−0.006	0.018
radius matching	7586	5202	**0.009**	[0.006]	1.49	[0.006]	−0.002	0.020
kernel matching	7586	5202	**0.006**	.	.	[0.006]	−0.004	0.020

Note: Significance at the 1%, 5%, and 10% levels is denoted by ***, **, and *, respectively.

Table 5.3: Estimation of the effect of gatekeeping restriction on the probability of the choice of the initial provider, OLS on (1) the full sample and (2) the sub-sample of respondents with exogenous choice of health insurance plan

	PCP				SPEC			
	(1)		(2)		(1)		(2)	
privGK	**0.036*** **	[0.009]	**0.032***	[0.018]	**-0.030*** **	[0.007]	**-0.014**	[0.015]
male	0.062***	[0.009]	0.043**	[0.018]	−0.029***	[0.007]	−0.018	[0.015]
hispanic	0.088***	[0.014]	0.134***	[0.028]	0.001	[0.011]	−0.037*	[0.022]
black	0.020	[0.014]	0.054*	[0.029]	0.000	[0.012]	0.005	[0.025]
years of educ	−0.008***	[0.014]	−0.062**	[0.030]	0.007***	[0.001]	0.003	[0.003]
NE	−0.058***	[0.014]	−0.062**	[0.030]	0.067***	[0.012]	0.040	[0.026]
MW	0.004	[0.013]	−0.027	[0.028]	0.012	[0.010]	−0.007	[0.023]
S	0.001	[0.012]	−0.020	[0.026]	0.053***	[0.010]	0.040*	[0.022]
married	0.013	[0.010]	0.041**	[0.020]	0.003	[0.008]	−0.015	[0.014]
log(income)	0.010	[0.007]	−0.006	[0.016]	0.036***	[0.006]	0.068***	[0.014]
# of children	0.005	[0.004]	0.012	[0.008]	−0.004	[0.003]	0.000	[0.007]
healthy	0.020**	[0.009]	−0.004	[0.018]	−0.016**	[0.008]	0.006	[0.015]
neoplasm	−0.131***	[0.034]	−0.070	[0.070]	0.348***	[0.037]	0.299***	[0.073]
diabetes	0.223***	[0.038]	0.354***	[0.055]	−0.131***	[0.026]	−0.136***	[0.047]
lipid metab.	0.170***	[0.029]	0.119*	[0.063]	−0.186***	[0.016]	−0.164***	[0.038]
blindness	−0.419***	[0.009]	−0.382***	[0.032]	0.207***	[0.035]	0.260***	[0.081]
hypertension	0.326***	[0.021]	0.357***	[0.039]	−0.170***	[0.015]	−0.193***	[0.027]
upper resp.	0.365***	[0.015]	0.331***	[0.033]	−0.200***	[0.009]	−0.223***	[0.019]
chron. pulm.	0.360***	[0.027]	0.318***	[0.054]	−0.219***	[0.011]	−0.240***	[0.020]
intestinal	0.253***	[0.031]	0.364***	[0.062]	−0.227***	[0.011]	−0.264***	[0.014]
urinary	−0.191***	[0.045]	−0.228***	[0.069]	−0.126	[0.033]	−0.123	[0.060]
pregnancy	−0.313***	[0.018]	−0.298***	[0.061]	0.366***	[0.027]	0.350***	[0.079]
disc disloc.	−0.136***	[0.019]	−0.182***	[0.037]	0.124***	[0.014]	−0.155***	[0.027]
sprain	−0.028	[0.024]	−0.091*	[0.048]	−0.125***	[0.016]	−0.139***	[0.034]
wounds	−0.171***	[0.035]	−0.128*	[0.072]	−0.198***	[0.016]	−0.186***	[0.035]
otherinj	−0.098***	[0.027]	−0.128***	[0.054]	−0.092***	[0.021]	−0.081***	[0.044]
joint	0.095***	[0.026]	0.086	[0.055]	0.018	[0.022]	0.038	[0.050]
connect. tissue	0.009	[0.024]	−0.060	[0.047]	−0.012	[0.021]	0.046	[0.046]
skin disorder	0.030	[0.021]	−0.002	[0.045]	0.147***	[0.021]	0.116***	[0.044]
constant	0.346***	[0.075]	0.511**	[0.169]	−0.233***	[0.064]	−0.431***	[0.148]
F-stat	191.49	[0]	25.66	[0]	63.53	[0]	22.72	[0]
R2	0.116		0.124		0.101		0.098	
N	12,789		2,973		12,789		2,973	

Note: Significance at the 1%, 5%, and 10% levels is denoted by ***, **, and *, respectively. Standard errors for coefficient estimates and p-value for F-statistics are reported in brackets.

Table 5.3: (continued from previous page)

	NONMD				HOSP				ER			
	(1)		(2)		(1)		(2)		(1)		(2)	
privGK	**-0.004**	[0.007]	**-0.019**	[0.014]	**-0.003**	[0.002]	**-0.005**	[0.005]	**0.001**	[0.005]	**0.007**	[0.011]
male	−0.062***	[0.007]	−0.066***	[0.014]	0.006**	[0.002]	0.003	[0.005]	0.023***	[0.006]	0.039**	[0.012]
hispanic	−0.080***	[0.010]	−0.076***	[0.019]	−0.003	[0.004]	0.008	[0.008]	−0.006	[0.009]	−0.029*	[0.018]
black	−0.082***	[0.010]	−0.084***	[0.020]	0.009*	[0.004]	0.000	[0.008]	0.054***	[0.010]	0.025	[0.020]
years of educ	0.006***	[0.001]	0.006**	[0.003]	−0.001	[0.000]	−0.000	[0.001]	−0.005***	[0.001]	−0.003	[0.002]
NE	−0.030***	[0.011]	−0.024	[0.003]	0.003	[0.004]	0.001	[0.007]	0.018**	[0.009]	0.045**	[0.019]
MW	−0.021**	[0.011]	−0.004	[0.023]	−0.002	[0.003]	−0.002	[0.006]	0.007	[0.008]	−0.040**	[0.017]
S	−0.061***	[0.010]	−0.058***	[0.020]	0.002	[0.003]	0.012*	[0.007]	0.005	[0.007]	0.026*	[0.015]
married	0.005	[0.008]	0.017	[0.016]	0.000	[0.003]	−0.004	[0.006]	−0.021***	[0.006]	−0.039***	[0.013]
log(income)	−0.012**	[0.006]	−0.023*	[0.013]	−0.002	[0.002]	−0.001	[0.004]	−0.032***	[0.005]	−0.030***	[0.010]
# of children	−0.011***	[0.003]	−0.019***	[0.006]	0.001	[0.001]	−0.001	[0.002]	0.009***	[0.003]	0.008	[0.006]
healthy	0.009	[0.007]	0.004	[0.014]	−0.012***	[0.003]	−0.012**	[0.005]	−0.001	[0.006]	0.006	[0.012]
neoplasm	−0.105 * **	[0.022]	−0.086*	[0.044]	−0.011	[0.008]	−0.021***	[0.004]	−0.101***	[0.009]	−0.122***	[0.010]
diabetes	0.055*	[0.032]	−0.077**	[0.036]	−0.022***	[0.007]	−0.027***	[0.005]	−0.124***	[0.008]	−0.115***	[0.011]
lipid metab.	0.156***	[0.028]	0.168***	[0.061]	−0.026**	[0.002]	−0.021***	[0.004]	−0.115***	[0.006]	−0.101***	[0.019]
blindness	0.341***	[0.036]	0.248***	[0.081]	−0.018***	[0.005]	−0.019***	[0.004]	−0.112***	[0.007]	−0.106***	[0.027]
hypertension	−0.058***	[0.015]	−0.055**	[0.028]	−0.014**	[0.006]	−0.006	[0.013]	−0.084***	[0.011]	−0.103***	[0.016]
upper resp.	−0.052***	[0.012]	−0.010	[0.027]	−0.019***	[0.003]	−0.017***	[0.006]	−0.093***	[0.007]	−0.081***	[0.016]
chron. pulm.	−0.073***	[0.020]	−0.022	[0.043]	−0.018***	[0.006]	−0.007	[0.015]	−0.050***	[0.017]	−0.049	[0.034]
intestinal	−0.063***	[0.021]	−0.078**	[0.039]	−0.007	[0.009]	−0.001	[0.022]	0.044*	[0.025]	−0.020	[0.047]
urinary	−0.087***	[0.027]	−0.114***	[0.031]	0.005	[0.019]	0.006	[0.032]	0.498***	[0.051]	0.459***	[0.088]
pregnancy	0.013	[0.022]	−0.049	[0.058]	0.037***	[0.0013]	0.065	[0.045]	−0.077***	[0.011]	−0.068**	[0.029]
disc disloc.	0.337***	[0.020]	0.416***	[0.040]	−0.018***	[0.004]	−0.010	[0.010]	−0.059***	[0.010]	−0.069***	[0.020]
sprain	0.037*	[0.020]	0.073*	[0.041]	−0.025***	[0.002]	−0.022***	[0.004]	0.140***	[0.021]	0.179***	[0.045]
wounds	−0.117***	[0.018]	−0.101***	[0.036]	−0.018***	[0.006]	−0.021***	[0.005]	0.504***	[0.038]	0.436***	[0.075]
otherinj	0.013	[0.022]	0.026	[0.044]	−0.025***	[0.002]	−0.022***	[0.004]	0.203***	[0.027]	0.205***	[0.054]
joint	0.035*	[0.021]	0.001	[0.040]	−0.020***	[0.004]	−0.011	[0.012]	−0.092***	[0.010]	−0.114***	[0.015]
connect. tissue	0.095***	[0.022]	0.095**	[0.044]	−0.018***	[0.004]	−0.003	[0.014]	−0.073***	[0.011]	−0.078***	[0.024]
skin disorder	−0.069***	[0.014]	−0.015	[0.031]	−0.017***	[0.004]	−0.022***	[0.004]	−0.091***	[0.008]	−0.077***	[0.020]
constant	0.312***	[0.060]	0.431***	[0.134]	0.059***	[0.023]	0.039	[0.048]	0.516***	[0.051]	0.449***	[0.106]
F	33.58	[0]	10.96	[0]	7.46	[0]	1.73	[0.01]	45.42	[0]	13.04	[0]
R2	0.081		0.106		0.012		0.014		0.110		0.121	
N	12,789		2,973		12,789		2,973		12,789		2,973	

Note: Significance at the 1%, 5%, and 10% levels is denoted by ***, **, and *, respectively. Standard errors for coefficient estimates, and p-values for F-statistics are reported in brackets.

Table 5.4: Probability of further visit conditional on the initial provider (in percentage points) - summary table

	probPCP		probSPEC		probnonMD		probHIS		probER	
First event	gate	non-gate	gate	non-gate	gate	non-gate	gate	non-gate	gate	non-gate
PCP	29.15	27.48	12.59***	10.01***	12.28	11.99	1.67	1.13	2.10	1.75
specialist	7.28*	5.74*	52.56	51.65	14.07	15.41	6.91	7.39	1.73	1.97
nonMD	12.80	13.72	15.11	14.30	43.50	45.87	2.93	2.59	1.97	1.34
hospital	30.33	21.88	34.43	41.67	16.39	17.70	10.66	15.63	28.69**	42.71**
ER	21.72*	17.84*	19.72	21.73	13.22	14.49	8.97	9.72	8.15	9.89

Note: The differences between gate and non-gatekeeping enrollees significant at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

Table 5.5: Number of further visit conditional on the initial provider - summary table

	probPCP		probSPEC		probnonMD		probHIS		probER	
First event	gate	non-gate	gate	non-gate	gate	non-gate	gate	non-gate	gate	non-gate
PCP	0.615	0.600	0.318**	0.240**	0.392	0.402	0.019	0.014	0.025	0.019
specialist	0.137*	0.097*	2.09*	1.84*	0.609	0.801	0.075	0.075	0.021	0.021
nonMD	0.228	0.282	0.500	0.468	2.687	2.978	0.030	0.030	0.025	0.015
hospital	0.475	0.5	1.95	1.521	1.484	0.739	0.164	0.177	0.311**	0.5**
ER	0.440	0.401	0.609	0.760	0.911	0.912	0.092	0.122	0.099	0.129

Note: The differences between gate and non-gatekeeping enrollees significant at the 1%, 5%, and 10% levels are denoted by ***, **, and *, respectively.

Table 5.6: Probability of further visits at different types of providers - marginal effects of initial provider choice, given gatekeeping status, from logit estimation on the full sample of respondents.

The model is estimated using a full set of dummy variables indicating combined initial provider choice - gatekeeping status, with non-gate PCP as the baseline category. Marginal effects are evaluated at the mean of other covariates; standard errors are computed by delta method and presented in brackets. Coefficients by gatekeeping subgroups represent the marginal effect of being a gatekeeping enrollee for a given type of provider.

	PCP		SPEC		nonMD		HOSP		ER	
first contact	logit	logit + X[a)]	logit	logit + X[a)]	logit	logit + X[a)]	logit	logit + X[a)]	logit	logit + X[a)]
PCP (base)	0.275	0.254	0.100	0.102	0.120	0.116	0.013	0.009	0.018	0.017
- PCP - gate	−0.017	0.011	0.026***	0.025***	0.003	−0.001	0.003	0.004	0.002	0.009**
	[0.012]	[0.012]	[0.009]	[0.009]	[0.009]	[0.009]	[0.004]	[0.003]	[0.004]	[0.004]
SPEC	−0.217***	−0.201***	0.416***	0.355***	0.034***	0.013	0.060***	0.031***	0.001	0.001
	[0.008]	[0.009]	[0.020]	[0.019]	[0.013]	[0.012]	[0.012]	[0.008]	[0.005]	[0.005]
- SPEC - gate	0.015*	0.017*	0.009	0.013	−0.013	−0.012	−0.010	−0.006	−0.002	−0.004
	[0.009]	[0.010]	[0.019]	[0.018]	[0.013]	[0.012]	[0.012]	[0.007]	[0.005]	[0.005]
nonMD	−0.138***	−0.128***	0.043**	0.021*	0.339***	0.263***	0.012*	0.006	−0.004	−0.007**
	[0.018]	[0.013]	[0.020]	[0.012]	[0.021]	[0.019]	[0.008]	[0.005]	[0.005]	[0.005]
- nonMD - gate	−0.009	−0.008	0.008	0.009	−0.024**	−0.017	0.005	0.003	0.006	0.006
	[0.013]	[0.014]	[0.014]	[0.012]	[0.020]	[0.023]	[0.008]	[0.006]	[0.007]	[0.006]
HOSP	−0.056	−0.051	0.317***	0.246***	0.057	0.040	0.143***	0.066**	0.331***	0.342***
	[0.043]	[0.042]	[0.053]	[0.048]	[0.040]	[0.036]	[0.042]	[0.029]	[0.057]	[0.061]
- HOSP - gate	0.085	0.067	−0.072	−0.081	−0.013	0.000	−0.035	−0.024	−0.115	−0.12*
	[0.059]	[0.057]	[0.066]	[0.061]	[0.051]	[0.046]	[0.004]	[0.032]	[0.065]	[0.060]
ER	−0.096***	−0.078***	0.117***	0.099***	0.025	0.012	0.084***	0.074***	0.056***	0.065***
	[0.018]	[0.018]	[0.020]	[0.019]	[0.017]	[0.015]	[0.018]	[0.016]	[0.014]	[0.014]
- ER - gate	0.039*	0.030	−0.020	−0.019	−0.013	−0.015	−0.043*	−0.004	−0.017	−0.015
	[0.021]	[0.021]	[0.022]	[0.020]	[0.019]	[0.017]	[0.023]	[0.014]	[0.162]	[0.013]
Wald stat.	604	1130	1696	1860	1133	1584	259	823	522	625
pseudo R2	0.059	0.109	0.130	0.173	0.088	0.151	0.067	0.232	0.119	0.149

Note: Significance at the 1%, 5%, and 10% levels is denoted by ***, **, and *, respectively.

a) Vector of covariates X consists of age, sex, ethnicity, years of education, marital status, health status dummy, and dummies for illness types.

Table 5.7: Number of further visits - marginal effects of initial provider choice and gatekeeping status from negative binomial regression on full sample

The model is estimated using a full set of dummy variables indicating combined initial provider choice - gatekeeping status, with non-gate PCP as the baseline category. Marginal effects are evaluated at the mean of other covariates; standard errors are computed by delta method and presented in brackets. Coefficients by gatekeeping subgroups represent the marginal effect of being a gatekeeping enrollee for a given type of provider.

	PCP		SPEC		nonMD		HOSP		ER	
first contact	nbreg	nbreg + X[a)]	nbreg	nbreg + X[a)]	nbreg	nbreg + X[a)]	nbreg	nbreg + X[a)]	nbreg	nbreg + X[a)]
PCP (base)	0.602	0.486	0.240	0.153	0.319	0.282	0.014	0.007	0.019	0.011
- PCP - gate	0.013	0.023	0.077**	0.041**	−0.010	−0.011	0.004	0.005	0.006	0.004
	[0.044]	[0.043]	[0.027]	[0.018]	[0.067]	[0.051]	[0.004]	[0.004]	[0.005]	[0.005]
SPEC	−0.505***	−0.471***	1.597***	1.272***	0.399***	0.356***	0.061***	0.033***	0.002	−0.001
	[0.020]	[0.021]	[0.161]	[0.133]	[0.145]	[0.138]	[0.013]	[0.010]	[0.005]	[0.005]
- SPEC - gate	0.040*	0.025	0.250*	0.306**	−0.193	−0.163	−0.001	−0.002	−0.001	−0.003
	[0.023]	[0.023]	[0.150]	[0.136]	[0.156]	[0.154]	[0.013]	[0.007]	[0.006]	[0.005]
nonMD	−0.320***	−0.319***	0.228***	0.123**	2.576***	2.187***	0.015*	0.009	−0.004	−0.003
	[0.036]	[0.033]	[0.072]	[0.055]	[0.410]	[0.353]	[0.008]	[0.006]	[0.005]	[0.005]
- nonMD - gate	−0.054	−0.042	0.032	−0.006	−0.291	−0.113	0.001	0.001	0.010	0.009
	[0.036]	[0.033]	[0.076]	[0.056]	[0.324]	[0.326]	[0.008]	[0.006]	[0.007]	[0.008]
HOSP	−0.102	−0.041	1.281***	0.891***	0.338	0.328	0.163***	0.069**	0.481***	0.413***
	[0.160]	[0.190]	[0.361]	[0.228]	[0.298]	[0.280]	[0.052]	[0.028]	[0.094]	[0.087]
- HOSP - gate	−0.025	−0.091	0.430	0.279	0.744	0.809	−0.013	−0.007	−0.189**	−0.193**
	[0.172]	[0.205]	[0.645]	[0.397]	[0.795]	[0.807]	[0.073]	[0.040]	[0.085]	[0.081]
ER	−0.201***	−0.147**	0.519***	0.424***	0.508**	0.325*	0.108***	0.093***	0.110***	0.081***
	[0.063]	[0.066]	[0.143]	[0.119]	[0.208]	[0.176]	[0.027]	[0.024]	[0.025]	[0.021]
- ER - gate	0.039	0.017	−0.151	−0.102	−0.001	0.033	−0.030	−0.022	−0.030	−0.029
	[0.078]	[0.077]	[0.148]	[0.123]	[0.245]	[0.203]	[0.022]	[0.020]	[0.023]	[0.018]
Pseudo-likelihood	−9515	−9191	−12147	−11659	−11592	−11142	−2212	−1906	−1849	−1174
Wald stat.	340	1046	930	64915	453	1384	237	72451	601	67490

Note: Significance at the 1%, 5%, and 10% levels is denoted by ***, **, and *, respectively.

a) Vector of covariates X consists of age, sex, ethnicity, years of education, marital status, health status dummy, and dummies for illness types.

Table 5.8: Probability of hospitalisation - comparison of PCP-referred and self-referred respondents for (1) the full sample and (2) the sub-sample of people without choice of insurance

	(1) logit + X[a)]		(2) logit + X[a)]	
PCP non-gate	0.063		0.053	
SPEC non-gate	-0.011	[0.018]	-0.006	[0.039]
PCP gate	0.030	[0.025]	0.046	[0.065]
SPEC gate	-0.024	[0.018]	-0.045	[0.032]
pseudo R2	0.226		0.139	
N	3,424		692	
Gate only:				
PCP referred - self-referred	0.052***	[0 .016]	0.094**	[0.045]
pseudo R2	0.217		0.156	
N	1,952		364	

Note: Significance at the 1%, 5%, and 10% levels is denoted by ***, **, and *, respectively.
a) Vector of covariates X consists of age, sex, ethnicity, years of education, marital status, health status dummy, and dummies for illness types.

Figures

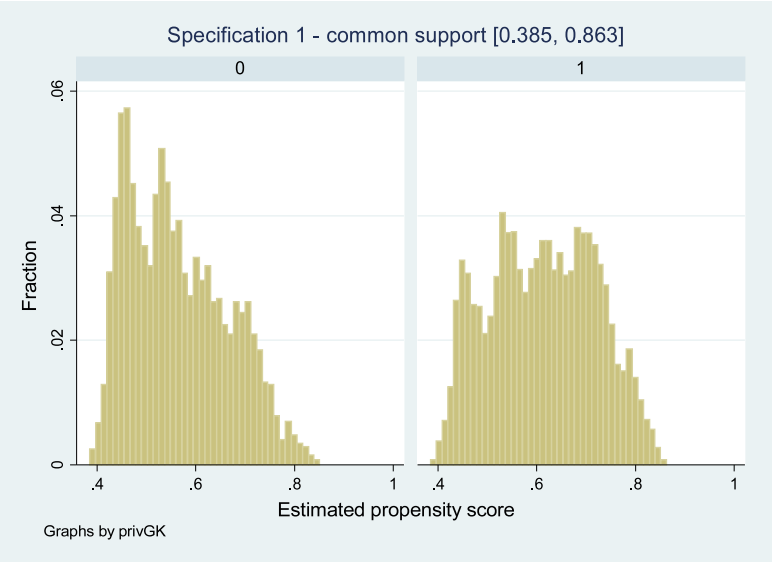

Figure 5.1: Histogram of propensity score from specification 1 by gatekeeping status

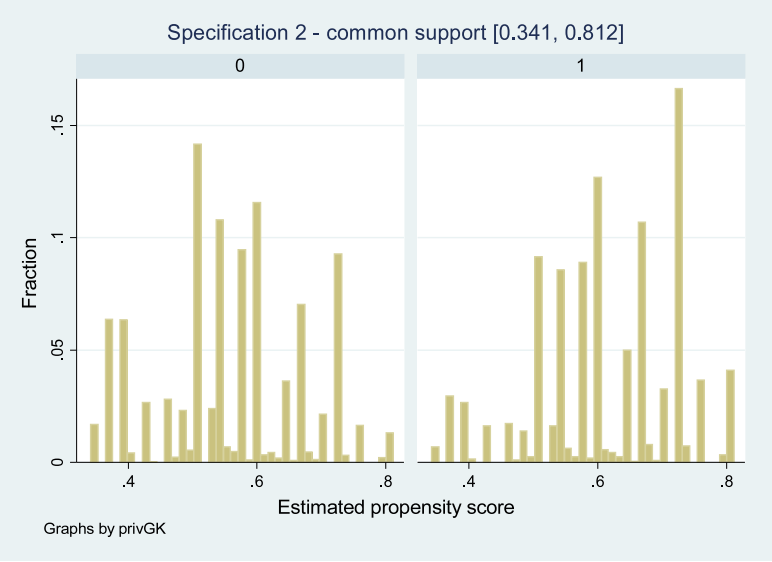

Figure 5.2: Histogram of propensity score from specification 2 by gatekeeping status

Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from the European Structural Fund (within the Operational Programme Prague Adaptability), the budget of the City of Prague, the Czech Republic's state budget and the following institutional grants:

- Center of Advanced Political Economy Research [Centrum pro pokročilá politicko-ekonomická studia], No. LC542, (2005-2009),
- Economic Aspects of EU and EMU Entry [Ekonomické aspekty vstupu do Evropské unie a Evropské měnové unie], No. AVOZ70850503, (2005-2010);
- Economic Impact of European Integration on the Czech Republic [Ekonomické dopady evropské integrace na ČR], No. MSM0021620846, (2005-2011);

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute ASCR, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: http://www.cerge-ei.cz

Editors: Directors of CERGE and EI
Managing editors: Deputy Directors for Research of CERGE and EI

ISSN 1211-3298
ISBN 978-80-7343-202-7 (Univerzita Karlova. Centrum pro ekonomický výzkum
a doktorské studium)
ISBN 978-80-7344-191-3 (Národohospodářský ústav AV ČR, v. v. i.)

CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
http://www.cerge-ei.cz