HEC Montréal

Affiliée à l'Université de Montréal

Trois essais en économie de la santé sur le comportement des médecins

par

Koffi Ahoto Kpelitse

Institut d'économie Appliquée
HEC Montréal

Thèse présentée à la Faculté des études supérieures et postdoctorales en vue de l'obtention du grade de
Philosophiae Doctor (Ph.D.)
en économie appliquée

Montréal, Québec, Canada
Décembre 2012

Université de Montréal

Faculté des études supérieures et postdoctorales

Cette thèse intitulée :

Trois essais en économie de la santé sur le comportement des médecins

présentée par

Koffi Ahoto Kpelitse

a été évaluée par un jury composé des personnes suivantes :

Président -rapporteur : Robert Gagné
Institut d'économie appliquée, HEC Montréal

Coprésident du comité de surveillance : Pierre Thomas Léger
Institut d'économie appliquée, HEC Montréal

Coprésidente du comité de surveillance : Marie Allard
Institut d'économie appliquée, HEC Montréal

Membre du jury : Roberk Clark
Institut d'économie appliquée, HECMontréal

Examinatrice externe : Rose Anne Devlin
Economics Department, University of Ottawa

Représentant du doyen de la FES : Eric Brunelle
Service de l'enseignement du management, HEC Montréal

Thèse acceptée le : 19 mars 2013

Sommaire

Cette thèse, constituée de trois essais, examine différents thèmes en économie de la santé sur le comportement des médecins à travers des modèles théoriques et empiriques.

Dans le premier essai, nous examinons le rôle de la concurrence entre les médecins dans un environnement où un patient qui décide de changer de médecin va à la recherche d'un nouveau médecin. De façon plus spécifique, nous présentons un modèle qui examine la relation entre le patient et le médecin et où les médecins sont caractérisés par le type de traitement qu'ils recommandent. Dans le modèle, après avoir reçus les soins fournis par son médecin, le patient observe son état de santé et décide soit de rester avec son médecin actuel ou soit de chercher un nouveau médecin en vue des soins futurs. Nos résultats montrent que des patients continuent de recevoir des soins de médecins qu'ils auraient dû quitter alors que d'autres patients quittent des médecins qu'ils n'auraient pas dû laisser. Nos résultats montrent également que la menace d'un patient de changer de médecin s'il ne reçoit pas des soins adéquats n'incite pas nécessairement son médecin actuel à lui fournir le traitement approprié. Par conséquent, certains patients peuvent recevoir trop de soins alors que d'autres en reçoivent moins.

Dans le second essai, nous examinons dans quelle mesure une combinaison des mécanismes de rémunération des médecins et des mécanismes de responsabilité pour faute médicale permet une offre efficace des soins de santé. À cette fin, nous déterminons de façon simultanée le niveau optimal de rémunération pour le médecin ainsi que le niveau de compensation optimale pour le patient étant donné que le médecin a commis une faute médicale. Notre modèle fournit des prédictions théoriques qui sont conformes avec ce que nous observons dans le monde réel : (i) les frais encourus par le patient pour poursuivre le médecin découragent la poursuite et (ii) le montant de la compensation versée au patient est proportionnel au dommage subi. Les résultats suggèrent également que toute politique permettant de réduire les coûts supportés par les patients en cas de poursuite pour faute médicale devrait permettre d'augmenter leur niveau de bien-être.

Le troisième essai examine dans quelle mesure les caractéristiques propres aux médecins influencent les variations du taux de césarienne à l'intérieur des hôpitaux canadiens. En utilisant des données institutionnelles où les patients sont plus susceptibles d'être associés de façon aléatoire aux obstétriciens, nous trouvons que les facteurs spécifiques aux médecins sont déterminants dans la décision de choisir une césarienne plutôt qu'un mode d'accouchement vaginal. Les résultats montrent également qu'à l'intérieur d'un même hôpital, la probabilité qu'une mère accouche par césarienne varie considérablement d'un médecin à un autre.

Mots clés : Système de rémunération à forfait, théorie du “search”, réseau social, responsabilité pour faute médicale, médecine défensive, césarienne, admission d'urgence, facteurs spécifiques aux médecins.

Summary

This thesis consists of three essays that examine different topics in health economics on physician behaviour in both theoretical and empirical models.

In the first essay we examine the role of competition amongst providers in a prospective payment setting using a patient's search framework. More specifically, we present a model which examines the patient-physician relationship where physicians are characterized by their treatment recommendation. Patients, following their health care consumption, observe their post-treatment outcome and decide either to continue to seek care from their current physician or search for an alternative one for future care. We find that some patients may stay with their physicians even if they do not receive the appropriate care while others may leave good physicians for poorer ones. The results also show that the fear of loosing a patient in a competitive market does not systematically induce the physician to treat appropriately the patient. As a result, the model generates equilibria with both under and over-provision of care.

In the second essay we analyze how medical-malpractice liability mechanisms combined with physician payment schemes may contribute to efficiency in the health-services market. To this end, we derive simultaneously optimal payment contracts for the physician and optimal compensation for the patient given that he is injured by the physician's action. The model predicts findings that are consistent with real world observations. It predicts that : (i) the fees incurred by the patient decrease the suing probability and (ii) the patient's compensation is proportionate to the damage. The results suggest that policies that may reduce malpractice lawsuit costs might be welfare improving.

The third essay analyzes whether and to what extent the variations in c-section rates are due to physician factors. Using institutional level data where patients are more likely to be randomly assigned to physicians, we find that physician factors are an important element in the c-section decision. The results show that, in the same hospital there is an important

variation across physicians in c-section rates.

Key words : Capitation, search theory, social network, medical malpractice, defensive medicine, c-section, emergent admission, physician factors.

Table des matières

Liste des tableaux

Liste des figures

Liste des annexes

À tous ceux et celles qui ont cru en moi.

Remerciements

Je tiens à adresser mes remerciements les plus sincères et les plus profonds à mes directeurs de thèse Marie Allard et Pierre Thomas Léger pour leur encadrement, leur grande disponibilité, leur patience ainsi que leur soutien moral et financier, qui m'ont permis de réaliser ce travail. J'ai beaucoup appris à leurs côtés durant tout ce cheminement doctoral et je leur adresse ici l'expression de ma grande reconnaissance.

Je remercie les professeurs Robert Gagné, Rose Anne Devlin, Robert Clark et Eric Brunelle d'avoir accepté de faire partie de mon jury de thèse. Qu'ils trouvent ici l'expression de mon profond respect et gratitude.

Je suis reconnaissant envers les professeurs Pierre Lasserre et Nicolas Sahuguet pour leurs commentaires et suggestions. Je remercie également l'Institut d'économie appliquée à HEC Montréal pour toute aide apportée durant toutes ces années.

Je remercie chaleureusement Marcus Loreti de l'Institut canadien d'information sur la santé pour toutes les discussions sur la base de données. Je remercie également Mohamed Jabir du Laced pour son aide et sa disponibilité.

Je dis un grand merci à mes fidèles compagnons de doctorat Ali Fakih, Ondelansek Kay, Hasina Rasata et Douré Grekou.

J'accorde un remerciement spécial à mes collègues de doctorat Fabienne Gouba, Michel-Yevenunye Keoula, Cédric Okou pour leur aide très appréciée. Je remercie également Tete Komla pour ses encouragements. Mes remerciements à mes amis Maurice, Damigou, Arlette, Thierry, Benis pour leur soutien.

Je termine en accordant un remerciement à mes parents, ma famille et à tous ceux qui m'ont soutenu.

Introduction générale

Au cours des dernières décennies, les dépenses en soins de santé n'ont cessé d'augmenter de façon croissante dans la plus part des pays développés et notamment dans ceux de l'Organisation de coopération et de développement économiques (OCDE). Selon les données de l'OCDE, en 2009 par exemple, les pays membres ont consacré en moyenne 9,6% de leur produit intérieur brut (PIB) aux dépenses de santé, avec en tête de liste les États-Unis qui y ont consacré 17,4%. Au Canada, pour la même année, selon l'Institut canadien d'information sur la santé (ICIS), les dépenses de santé du secteur public ont atteint un sommet, soit 8,5 % PIB alors qu'à la fin des années 1970, ce ratio ne dépassait qu' à peine les 5 %.

Parmi les facteurs souvent invoqués pour expliquer cette hausse des dépenses en soins de santé, on retrouve les facteurs démographiques (notamment le vieillissement de la population), le progrès technique et les différents mécanismes de rémunération médecins. Au Canada, selon l'ICIS, ce sont les dépenses consacrées aux médecins qui figurent parmi les facteurs ayant connu la plus forte croissance ces dernières années dans le secteur de la santé, atteignant 6,8 % par an de 1998 à 2008 et plus de la moitié de cette croissance est liée à la hausse des rémunérations des médecins. [1]

Dans un contexte d'asymétrie de l'information entre le médecin et le patient, les mécanismes de rémunération comme le paiement à l'acte peut inciter le médecin à induire la demande et par conséquent provoquer une hausse des coûts des soins de santé.

Cette thèse est composée de trois essais qui examinent différents thèmes en économie de la santé sur le plan théorique et empirique. De façon spécifique, nous nous intéressons aux comportements des médecins dans divers environnements. Étant donné que les décisions des médecins ont souvent des répercussions importantes sur les coûts et la qualité des

[1] La croissance des dépenses liées au viellissement de la population est très modérée et est d'à peu près 0,8% annuellement.

soins de santé, une analyse des comportements tant stratégiques que non stratégiques des médecins s'avère nécessaire et voire indispensable. En effet, elle devrait permettre de mieux comprendre les incitations des médecins et de déterminer les mécanismes ou les politiques à mettre en place pour atteindre l'efficacité dans le secteur des services de soins de santé. Dans les deux premiers essais, nous développons des modèles théoriques dans le but d'analyser les mécanismes (tant au niveau de la demande que de l'offre) pouvant permettre une offre et une consommation efficaces des soins de santé. Dans le troisième essai, nous utilisons des données canadiennes pour examiner les variations du taux de césarienne dans plusieurs hôpitaux et différentes provinces.

Le premier essai s'intéresse au rôle de la concurrence entre les médecins. Dans le but de contrôler l'augmentation croissante des coûts des soins de santé observée dans la plupart des pays développés au cours des dernières décennies, plusieurs pays ont adopté un système de rémunération à forfait (par capitation), i.e., un paiement fixe que le médecin reçoit pour chaque patient traité ou inscrit dans sa pratique. Malgré le fait que ce type de rémunération soit efficace pour freiner la hausse des coûts, les modèles théoriques prédisent qu'il peut encourager (du moins dans un cadre statique) les fournisseurs de services de soins de santé à réduire la qualité et la quantité de soins nécessaires. Cependant, ces modèles ne tiennent pas compte de l'impact de la concurrence entre les médecins. En effet, la menace d'un patient de changer de médecin si les soins reçus ne sont pas appropriés peut imposer une limite à la capacité des médecins de réduire les soins de santé.

Peu d'auteurs ont analysé le rôle de la concurrence entre les médecins dans le cadre d'un mécanisme de paiement donné. Ceux qui l'ont fait ont supposé qu'un patient qui décide de quitter son médecin est assigné à un autre médecin, soit de façon exogène (Dranove, (1988) ; Rochaix, (1989) ; Ma et McGuire, (1997)), ou soit de façon aléatoire (Allard, Léger et Rochaix, (2009)). Dans cet essai, nous examinons aussi le rôle de la concurrence entre les médecins mais en supposant qu'un patient qui décide de changer de médecin va plutôt à la recherche d'un nouveau médecin et consulte son réseau social pour s'informer sur la qualité de ce nouveau médecin. Par conséquent, la principale contribution de cet essai est d'incorporer la "théorie du search" dans les modèles traditionnels de concurrence. Nous pensons que cette formulation décrit de façon plus adéquate le comportement observé des patients, lorsqu'ils choisissent leur médecin ou qu'ils décident de changer de médecin. De plus, elle permet d'obtenir des résultats qui sont cohérents avec les faits stylisés.

Les résultats montrent que les patients commettent des erreurs de type I et II. En d'autres termes, des patients continuent de recevoir des soins de médecins qu'ils auraient dû quitter alors que d'autres patients quittent des médecins qu'ils n'auraient pas dû laisser. Par conséquent, des mesures visant à fournir de l'information sur la qualité des médecins pourraient augmenter le bien-être des patients. Cette instabilité de la relation entre les patients et les médecins est également présente dans Allard, Léger et Rochaix, (2009), mais leur modèle n'est pas en mesure de générer les deux types d'erreur de façon simultanée. Nos résultats montrent également que la menace d'un patient de changer de médecin s'il ne reçoit pas des soins adéquats n'incite pas nécessairement son médecin actuel à lui fournir le traitement approprié. Par conséquent, certains patients peuvent recevoir trop de soins alors que d'autres en reçoivent moins.

Le deuxième essai examine dans quelle mesure une combinaison des mécanismes de rémunération des médecins et des mécanismes de responsabilité pour faute médicale permet une offre efficace des soins de santé. Les décisions des médecins ont souvent des conséquences en termes de réduction des coûts des soins de santé et de l'amélioration de la qualité des soins. Ces deux derniers objectifs étant conflictuels, l'utilisation simultanée de deux mécanismes peut s'avérer plus efficace pour leur atteinte. Cependant, dans la littérature, peu de travaux ont étudié les effets de la combinaison de ces deux instruments ou mécanismes.

Les papiers qui ont examiné le rôle de ces deux mécanismes l'ont fait dans un environnement où l'interdépendance entre les deux mécanismes est ignorée et où, soit les mécanismes de responsabilité pour faute médicale (Gal-Or, (1999) ; Zeiler, (2008)), soit les mécanismes de rémunération (Léger, (2000) ; Arlen et MacLeod, (2005)) sont exogènes. Pourtant, un système de responsabilité pour faute médicale peut avoir des impacts qui dépendent des systèmes de rémunération. La contribution de cet essai réside dans la prise en compte de cette interdépendance en traitant de façon endogène les deux mécanismes.

Dans un premier temps, nous retrouvons les résultats de Léger, (2000) et Zeiler, (2008) qui montrent qu'à l'équilibre, le médecin ne fournira jamais les soins appropriés avec certitude (avec probabilité égale à 1) ; et en réponse à cette stratégie, le patient ne va jamais poursuivre le médecin avec certitude. Par la suite, le fait de considérer des mécanismes de paiement flexibles nous permet de générer un équilibre dans lequel le bien-être du patient est relativement plus grand comparativement au système à forfait pur. Notre modèle fournit des prédictions théoriques qui sont conformes avec ce que nous observons dans le monde réel : (i) les frais encourus par le patient pour poursuivre le médecin découragent la poursuite et (ii) le montant de la compensation versée au patient est proportionnel au dommage subi. Les résultats suggèrent également que toute politique permettant de réduire les coûts supportés par les patients en cas de poursuite pour faute médicale devrait permettre d'augmenter leur niveau de bien-être.

Finalement, le troisième essai porte sur les variations du taux de césarienne à l'intérieur des hôpitaux canadiens. Plusieurs études ont montré qu'il existe des variations importantes au niveau des pratiques médicales dans plusieurs domaines des services de santé[2]. Ces variations peuvent entrainer une perte de bien-être si les fournisseurs de soins dévient de façon

[2] Parmi les plus récentes, on peut citer : Epstein et Nicholson, (2009) ; Bynum, Song et Fisher, (2010) et Tu et al., (2012).

significative du niveau efficace d'offre de soins de santé. Outre les mécanismes de rémunération, des facteurs comme les caractéristiques des médecins (Phelps, (2000) ; Grytten et Sorensen (2003) et Epstein et Nicholson, (2009)) et des hôpitaux ont souvent été cités comme les déterminants de ces variations. Au cours des dernières décennies, le taux d'accouchement par césarienne n'a cessé d'augmenter dans la plupart des pays développés. Au Canada, ce taux était d'environ 28% en 2008[3], soit à peu près le double du taux de 15% recommandé par l'Organisation Mondiale de la Santé en 1985. Dans cet essai, nous examinons dans quelle mesure les caractéristiques propres aux médecins influencent les variations du taux de césarienne au sein des hôpitaux.

Des études ont montré qu'il existe des différences significatives dans la propension et la capacité des obstétriciens à effectuer ou à recommander une césarienne. Cependant, certains de ces résultats peuvent être biaisés à cause des problèmes d'endogénéité. En effet, Grant, (2005) et Epstein et Nicholson, (2009) ne tiennent pas compte de façon explicite du problème de sélection des patients et/ou des médecins ni du fait que certains médecins traitent un nombre relativement élevé de patients avec des conditions médicales particulières. Même si Epstein, Ketcham et Nicholson, (2010) tiennent compte de ces problèmes d'endogénéité, en utilisant un échantillon où les patients sont plus susceptibles d'être associés de façon aléatoire aux obstétriciens, le problème de la sélection n'est pas complètement éliminé. La principale contribution de cet essai est d'utiliser les données sur les admissions d'urgence pour contrôler pour les problèmes d'endogénéité.

Les résultats montrent que les facteurs spécifiques aux médecins sont déterminants dans la décision de choisir une césarienne plutôt qu'un mode d'accouchement vaginal et que ces facteurs expliquent une part importante des variations observées dans les taux de césarienne. Les résultats montrent également qu'à l'intérieur d'un même hôpital, même en contrôlant

[3] Selon les rapports de l'institut canadien d'information sur la santé.

pour les conditions médicales des patients, la probabilité qu'une mère accouche par césarienne varie considérablement d'un médecin à un autre. Par conséquent, les politiques qui encouragent une standardisation des procédures à l'intérieur des hôpitaux pourraient augmenter le bien-être social.

Essay 1
Competition in the Physicians' Market : A Search Theoretic Approach

Abstract

In this paper, I examine whether and to what extent a patient's search for a competing physician may encourage efficiency in a cost-containment prospective-payment setting. More specifically, I build a two-period model where patients, after health care consumption, evaluate their physicians' quality, and may decide (following a new health shock) to stay with him or to search for an outside physician and potentially switch. I find that the patient's switching decision depends on his level of ignorance about both the illness severity and the type of the alternative physician. Depending on the extent of these elements, patients make type I and II errors. That is, some patients may stay with their physicians even if they do not receive the appropriate care while others may leave good physicians for poorer ones. By allowing physicians to act strategically, I find under some assumptions that the treatment provided depends on the physician's beliefs about the patient's level of ignorance. As a result, the model generates equilibria with both under and over-provision of health care.

1 Introduction

The presence of information asymmetry, health insurance and certain physician payment mechanisms can lead to inefficient consumption of healthcare. In such environments, it has been shown that only contracts or payments based on health states are efficient (Arrow, (1963)). However, because health states are unverifiable (or, at least are costly to identify), such contracts are infeasible. In response to this infeasibility, several solutions have been proposed in order to reach a second-best outcome.

Among these solutions are demand-side and supply-side incentives. On the supply side, prospective payments such as capitation (a payment system in which physicians receive a fixed payment for each patient they enlist into their practice) may be an effective way to deal with excessive provision of health-care services and rising costs.[4] Such payments are, however, often associated (at least theoretically) with the under provision of care.[5] This theoretical prediction, however, relies in part on the questionnable assumption of limited competition between providers. More specifically, the fact that patients can leave their current physician for an alternative one if they are unsatisfied with their care, may reduce physicians' willingness or ability to underprovide medical care. Although a few papers in the literature consider the role of competition in a payment framework, they do so in an environment where patients' outside options are either exogenously given (Dranove, (1988) ; Rochaix, (1989) ; Ma and McGuire, (1997)), or relies on random matching in a model of endogeneous competition (Allard, Léger and Rochaix, henceforth ALR, (2009))[6]. In this

[4] Because physicians are often the key decision makers in most of health care resources' utilization, it has been argued that supply-side cost control mechanisms such as capitation are likely to be more efficient than demand-side cost controls for several reasons (Ellis and McGuire (1993)).

[5] Because providing more care to the patient is costly, physicians have an incentive to shrink on care.

[6] Other mechanisms such as monitoring by third parties (Léger, (2000)) and liability from medical malpractice (Danzon, (2000)) may be used to achieve efficient delivery of health care under such a prospective system. It is also important to note that the presence of altruism or ethical concerns may limit the incentive for physician's desire to shrink on care even in the absence of monitoring or regulation (Ma and McGuire, (1997) ; Jack, (2005) ; ALR, (2009)).

paper, I also examine the role of competition amongst providers in a prospective payment setting but within a search framework. More specifically, I examine whether and to what extent a patient's search for a competing physician may encourage the efficient provision and consumption of care in a cost-containment prospective-payment setting. In fact, as it will be shown, integrating search into the normal framework allows us to explain some stylized facts, which cannot be described in the traditional models without additional assumptions.

I present a model which examines the patient-physician relationship where physicians are characterized by their treatment recommendations. Patients, following their health care consumption, observe their post-treatment outcome and decide either to continue to seek care from their current physician or search for an alternative one for future care.

Unlike previous models, this paper focuses on a framework that includes the elements of search theory. The utilization of this framework is motivated by several factors. First, it is consistent with the observed behaviour of patients when choosing physicians or deciding for alternative physicians (see for example Gourash, (1978) ; Hoerger and Howard, (1995) ; Kaiser Family Foundation/Agency for Health Care Research and Quality, (2000)). Second, the patient-physician relationship is unstable and empirical evidence shows that 4 to 11% of patients switch physicians annually (see for example Sobero et al. (2003)). As argued by Ma and Choné (2011), patients generally search for the right physicians and as a consequence, they suggest that the models which examine competition among physicians for patients must integrate the matching process between patients and physicians.[7] Third, patients do not usually sign contracts with their physicians or pay them directly and thus the traditional monitoring mechanisms of the principal-agent model, which define incentive constraints, cannot be used to overcome the asymmetric information problem between the patient and the physician (Rochaix, 1989). As a consequence, the patient's search among physicians

[7] A right physician refers to a physician who will provide the clinically appropriate treatment given the patient's illness severity.

may serve as an ex ante monitoring device which may encourage the physician to provide appropriate care.

The model predicts that the patient's switching decision depends on his level of ignorance about his illness severity as well as on the information obtained about an alternative physician. Depending on the extent of these elements, some patients may stay with their physicians even if they do not receive the appropriate care while others may leave good physicians for poorer ones. In other words, patients in the model make both type I and II errors. Thus the model's predictions are consistent with the medical evidence (Rossister et al., (1989) ; Schlessinger et al., (1999)) which concludes that some patients continue to seek care from a physician they believe to be incompetent.

By allowing physicians to act strategically, I find that the fear of loosing a patient in a competitive market does not systematically induce the physician to treat appropriately the patient. More specifically, the treatment provided depends on the physician's beliefs about his or her patient's type. As a result, the model generates equilibria with both under and over-provision of care.

As mentioned briefly before, this paper is related to work on non-price competition in the physicians' market. In Dranove (1988), physicians compete on the level of "inducement of the demand", where the patient's willingness to leave a physician who is too aggressive in his treatment recommendations can provide a disincentive for inducement behaviour. Dranove (1988) however remains silent on how patients choose their alternative providers. Similarly, in Rochaix (1989) the patient's threat to seek a second opinion serves as a monitoring device for physicians (i.e., it induces physicians to choose a level of treatment which is close to what their patients would choose if they were well informed). As in Dranove (1988) the patient's outside option is exogenously given. Thus, both Dranove (1988) and Rochaix (1989) assume that the patient's decision to accept treatment is based on the gap between his or her expectation about the illness severity and the physician's recommendation. However, the assumption that patients can evaluate their physician's quality prior to being treated is somewhat questionnable. I address these two limitations by (i) endogenizing the competition between physicians, and (ii) allowing patients to judge their physician's quality only after receiving care from him or her.

ALR (2009) study the repeated interactions between patients and physicians in a theoretical model where physicians are differentiated by an unobserved individual-specific ethical constraint which specifies the minimum illness-specific amount of effort a physician is willing to provide to his or her patient. Although they endogenize the patient's outside option, they assume that a patient who decides to leave his current physician is randomly matched to a new physician. This assumption may not be consistent with the stylized fact that patients rely on information or recommendations from their family members, friends or co-workers in their needs of medical goods. That is, they collect information on particular outside physicians before making a switch. By ignoring such search behaviour, ALR (2009) cannot generate both type I and II errors simultaneously in equilibrium. That is, they cannot generate an equilibrium where some patients leave their current physician for a worse one while others fail to leave bad ones for higher quality ones simultaneously.

The remainder of this paper is organized as follows. Section 2 presents the model. In section 3, I solve for patients' search behaviour. In section 4, I look for the role of competition between physicians. The last section contains some concluding remarks and policy implications. All proofs are in the Appendix B.

2 The Model

In this section, I present a model of the relationship between patients, physicians and an insurance company. In the model, the insurer who is assumed to operate in a perfectly competitive market, collects premiums and pays physicians on a capitated basis.[8] I also assume that the insurer signs a contract with physicians and patients before the patient's illness is revealed.

In the first period, if an individual becomes ill, he seeks care from the physician. The physician observes the patient's illness severity and provides a treatment which is assumed to be observable by both the patient and insurer. I assume that there is an appropriate (exogenous) level of treatment which is associated with every illness severity, and that this is public information.[9] It is important to note that in the model, the appropriate treatment refers to the illness-specific treatment that a fully informed patient would choose. After healthcare services are consumed, the post-treatment health is revealed to both the patient and the doctor. In the second period, if the patient becomes ill again, he draws a new illness severity independently but from the same distribution as in the first. Next, the patient must decide whether to continue to seek care from the same physician or to search for a new one.[10] Furthermore, it is assumed that if a patient does not receive a new health shock, there is no need for care, and consequently no search.[11] This model is particularly well suited to the relationship between patients with chronical illnesses and their specialists.[12]

[8] I assume that the patient has full insurance, thus there is no co-payments.

[9] The major implication of this assumption is that patients and physicians have the same health treatment function. This is reasonable because the appropriate treatment is exogeneous and represents the clinically appropriate treatment.

[10] Even if the new health shock is assumed to be drawn from the same illness distribution, the level of severity may differ. That is, I do not impose any restriction on the relationship between the first and the second health shock.

[11] This may not be true in a multi-period setting because a risk averse patient may search as a form of precaution.

[12] Although the severity of chronic illnesses may be correlated over time, the assumption of their indepen-

I describe in the following sections the timing of the model and the agents' preferences.

2.1 The Timing

The timing of the game is as follows :

Period 1

Step 1

The insurance company offers contracts to the physicians and patients. These contracts specify the capitation payment P for the provider and the premium α for the insured individual. The premium α is assumed to be actuarially fair.

Step 2

With probability π, the patient receives a health shock and requires treatment, and with probability $(1-\pi)$, he is healthy. If a patient is ill, he or she draws a particular θ from a known distribution function $F(\theta)$ with support $[\theta_L, \theta_H]$ and where θ_L corresponds to the lowest level of severity and θ_H the highest level. I also assume that the patient, given his symptoms, estimates his illness severity with error as $\theta^p = \theta + \mu$ and thus expects to receive the corresponding appropriate level of treatment $t^*(\theta^p)$. μ is a parameter which denotes the patient's type and is fixed over time.[13] As a result, μ can be thought of as the level of ignorance of each patient about his true health condition : if $\mu > 0$ then the patient is an over-estimator of his illness severity while $\mu < 0$ refers to an under-estimator. Thus μ can be thought of as an individual-specific effect. I assume that the mean of μ is equal to zero in the population of the patients, and that patients' types are distributed according to a distribution function $T(\mu)$.[14]

---

dency is only for simplification purposes. This simplification makes the mathematics more tractable.

[13] That is, I assume that there is no learning. This assumption seems reasonable in a two-period model but not be true in a multi-period setting.

[14] The assumption that $E(\mu) = 0$ seems reasonable in our context. Even if some patients, in reality, are hypochondriacs, their over-estimation may be compensated by those who under-estimate their illness severity.

Step 3

I assume that the physician, given his medical knowledge and clinical information, perfectly observes the patient's illness severity θ. After observing θ, the physician provides a level of treatment, which is denoted by $t(\theta)$. This level of treatment may be different from the appropriate treatment given θ (i.e., $t^*(\theta)$). More specifically, I assume that $t(\theta) = \delta t^*(\theta)$ where δ is a fixed and positive parameter which captures the physicians' unobserved heterogeneity. I assume that $0 < \delta \leq \bar{\delta}$, where $\bar{\delta}$ designs the upper bound and may be greater than one.[15] The physician types are distributed according to a known distribution function $G(\delta)$ and $g(\delta)$ denotes the density function.

Step 4

Given that the patient can observe $t(\theta)$, he can compare his expected treatment $t^*(\theta^p)$ to his current treatment $t(\theta)$ and form an estimate of his physician's type.[16]

Period 2

Step 5

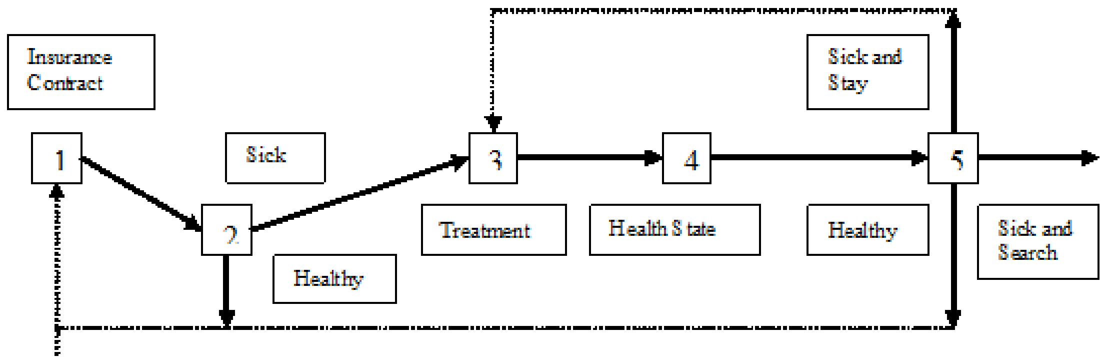
If the patient receives a second health shock from the same illness distribution, he may decide either to stay with the same physician or to search for another one.[17] If the patient is healthy, then the game ends. Following the new health shock, if the patient decides to search for another provider, he must first choose how intensively to search. Once a contact with a new physician is made, the patient collects informations about the type of that particular physician. Following the signal received about that physician's type, the patient must then decide whether to switch or to stay with his current physician.

[15] Since δ can be greater than 1, I allow overprovision of health care in the model but I assume that it affects only the health-care costs. That is, the overprovision is not harmful for the patient's health.

[16] I assume that the treatment-outcome relationship is certain.

[17] Search for another physician is costly and since the second period's health care consumption is conditional on receiving a new health shock, patients do not automatically search after the first period's treatment.

Figure 1.1 : The Timing



2.2 The Patients

As discussed above, patients have only one kind of disease but are heterogeneous relative to the level of severity of their health shocks. As noted before, each patient is characterized by a μ parameter which is considered as his or her type. The patient's utility is assumed to depend on two elements : health state h and income available for consumption goods x where utility is increasing and concave in both elements. The patient's health state depends negatively on the illness severity and positively on the quantity of treatment received.

Formally, the patient's per-period expected utility is :

$$U^p = (1-\pi)U(x, h(0,0)) + \pi \int_\theta U(x, h(\theta, \delta t^*(\theta))dF(\theta) \tag{1}$$

where $x = y - \alpha$ and y is the individual's income ; $h(0,0)$ denotes the patient's health in the absence of illness.

Equation (1) states that the patient's per-period expected utility also depends on whether or not he receives a health shock. If he receives a health shock, then his utility depends on both the severity of the shock and the type of his physician.

2.3 The Physician

As noted above, physicians differ in terms of their treatment recommendation and I refer to this as the physician's type δ. The type parameter δ is interpreted as the proportion of the appropriate treatment provided by the physician. Thus $\delta = 1$ corresponds to the case of the physician who provides the appropriate treatment. A $\delta < 1$ ($\delta > 1$) corresponds to a physician who under-treats (over-treats) the patient. The heterogeneity in physicians' market can be seen as : for the same level of severity, patients may receive different treatments since the choice of the treatment depends on the physician's type.

2.4 The Insurance Provider

The insurance market is assumed to be perfectly competitive where profit maximizing insurers collect premiums and pay physicians on a capitated basis. I also assume that insurers do not restraint patients in their choice of health-care providers. Since the insurer can observe ex-ante neither physician type nor patient's illness severity, writing contracts based on physician type or illness severity is infeasible (the first best is unachievable).

The equilibrium parameters are found as follows : (i) I solve for the patient's optimal response given the physician's treatment decision ; and (ii) I resolve the physician's problem given the search behaviour of the patient.

3 The Patient's Search Behaviour

In this section, I model the patient's switching and search behaviour in an imperfect information (about the illness severity) setting. As noted before, I assume that the patient cannot perfectly observe his illness severity θ but estimates it as θ^p. Because of the difference between θ^p and θ the patient cannot attribute all the difference between his expected post-treatment health and his current post-treatment health to the physician (and he is aware of this). In other words, some of the difference may be attributed to the patient's incorrect estimation of his illness severity. In this setting, the patient compares the expected treatment $t^*(\theta^p)$ to his current treatment $t(\theta)$ to decide whether or not to search.

Recall that $\theta^p = \theta + \mu$ where the patient's type μ is drawn from a known distribution $T(\mu)$. Hence, the patient's expected and realized treatment can be rewritten as $t^*(\theta + \mu)$ and $\delta t^*(\theta)$, respectively, and the corresponding expected post-treatment health and realized post-treatment health as $h(\theta+\mu, t^*(\theta+\mu))$ and $h(\theta, \delta t^*(\theta))$, respectively. As assumed above, patients evaluate their physicians' types by comparing their expected treatment to their realized treatment, i.e., they compare $t^*(\theta + \mu)$ to $\delta t^*(\theta)$. From this comparison the patient can estimate his current physician's type. Let denotes by δ^{EV} the patient's evaluation of his physician's type. It follows that $\delta^{EV} = \frac{\delta t^*(\theta)}{t^*(\theta+\mu)}$.

Depending on the values of μ, three possible cases may arise. (1) if $\mu = 0$, then $\delta^{EV} = \delta$, i.e., when the patient estimates perfectly the illness severity, he perfectly infers his physician's type. (2) if $\mu < 0$, then $t^*(\theta) > t^*(\theta + \mu)$ or equivalently $\frac{t^*(\theta)}{t^*(\theta+\mu)} > 1$ and hence $\delta^{EV} > \delta$; i.e., when the patient under-estimates the illness severity, he over-evaluates his physician's type.[18] (3) if $\mu > 0$, using the same rationale as above, we have $\delta^{EV} < \delta$, i.e., when the patient over-estimates the illness severity, he under-evaluates his physician's type.

It is important to note that the above results follow from the fact that I have assumed that : (i) the patient observes the treatment and (ii) there is an appropriate level of treatment which is associated with every illness severity, and that this is public information. In fact, without these assumptions, an over-estimator of the illness severity may over-evaluate his physician's type. However, these assumptions do not imply that the physician will always provide the appropriate treatment. In fact (as it will be shown later), even if the patient is perfectly informed, an inappropriate treatment may occur because the probability that the

[18] I assume that t^* is increasing in θ.

patient will leave (his current physician) given his search behaviour is less than one.

For a given θ and δ, the patient's ex post-treatment's utility level is : $U^p = U\left[x, h(\theta, \delta t^*(\theta))\right]$. The patient knows that if he receives another health shock, his current physician will treat him with the same proportion of the appropriate treatment ; that is, the physician is assumed to behave in the same way. Thus, the patient knows that if he stays with the current physician, he will get a level of utility equal to $U(x, h(\theta, \delta t^*(\theta)))$.[19] By deciding to switch, the patient searches first for an alternative physician. If a new physician is found, then given the type of that specific physician, the patient decides whether or not to switch. The matching technology is given in more detail below.

3.1 The Patient's Search Strategy

I start by describing the patient's search strategy. Much of my description of patient's search behaviour is borrowed from the abundant literature on assignment, search and matching.[20] Although search theory is extensively developed, I restrict my attention to a few papers which are particularly close to the present work.

Using the standard search framework (Mortensen(1986), Burdett & Coles(1999) and Wolinsky (1987)), I present a single-sided search model where patients search for a "better physician". The matching technology is described as follows : a patient contacts a physician according to a Poisson process with parameter λ. So the probability that an individual i will meet a physician during short interval Δ is $\lambda_i \Delta$. This probability can be raised by the patient, but raising the meeting probability is costly. Mortensen (1982) terms λ_i as person-specific search intensity. I follow Mortensen (1982) by assuming that the probability of a new match formation is determined by the search intensity chosen by the patient. For the

[19] The illness severity can differ, so θ is not necessary the same as in the first period.

[20] Search theory is used to explain the rate of unemployment, wage dispersion among homogeneous workers (in labor market) ; price dispersion of homogeneous goods (in product market) ; the matching process between men and women (in marriage market) ;...

purpose of this paper, I allow patients to recall a past observation, that is, if the search process is unfruitful, the patient can return to his previous physician.

3.2 The Equilibrium Analysis

Ex ante, the patient has no information about his physician's type but ex post he can form beliefs about his physician's type by comparing $t^*(\theta + \mu)$ to $\delta t^*(\theta)$. Henceforth, to simplify notation, I rewrite the patient's post-treatment health as a function of his illness severity and the type of physician he patronizes, i.e., $h = h(\theta, \delta)$.

Let V^p denote the expected present value from search and δ^c the true type of the current physician (and thus $U^p = U(x, h(\theta, \delta^c))$ the patient's current level of utility). As a consequence, a risk-neutral patient would stay with his current physician if and only if : $U(x, h(\theta, \delta^c)) \geq V^p$.

Next, I compare : (i) for patients who do not switch, the treatment they would receive if they had left with that actually received and, (ii) for those who switched, the treatment received with that which they would have received if they had stayed with their physician. More specifically, I investigate whether the search process can help patients taking appropriate decisions.

3.2.1 The Case Where the Outside Physician's Type Is Perfectly Observable

In this section I assume that once a new physician is met, the patient perfectly and instantaneously observes physician type. I relax this assumption in the next section (3.2.2).

In order to maximize his expected utility from receiving care from the alternative physician, the patient chooses for every level of evaluation δ^{EV}, a search intensity λ^* and a critical physician type δ^* to maximize the following expression :

$$V^p = \max_{(\lambda^*,\delta^*)} \left\{ -\Delta c(\lambda) + \frac{1}{1+r\Delta} \left[\begin{array}{c} \lambda\Delta \left(\begin{array}{c} \Pr(\delta > \delta^*) \times E(V^p|\delta > \delta^*) + \\ \Pr(\delta \leq \delta^*) \times U^p \end{array} \right) \\ + \\ (1-\lambda\Delta)\, U^p + o(\Delta) \end{array} \right] \right\} \tag{2}$$

where r denotes the discount rate, $c(\lambda)$ the cost of search incurred by the patient during the time interval Δ with $c' > 0$, $c'' > 0$, $c(0) = c'(0) = 0$ and $o(\Delta)$ the probability that the patient meets more than one physician in the interval Δ, where $o(\Delta)$ tends to 0 when Δ approaches 0.

Equation (2) states that during the short time interval Δ, the patient is looking for a new physician and pays search cost $c(\lambda)$. With probability $\lambda\Delta$ he meets a physician. Once a physician contact is made, the type of this physician can be either acceptable (i.e., relatively high) or not to the patient. If the type is acceptable to the patient (i.e., if $\delta > \delta^*$) then the patient will switch and his expected utility will be V^p, otherwise he will stay and get U^p (i.e., if $\delta \leq \delta^*$).[21] Finally, with probability $(1 - \lambda\Delta)$ he does not meet any alternative physician and must stay with his current physician and get U^p.

Equation (2) can be rewritten as (for the Proof, see the Appendix A) :

$$V^p = \max_{(\lambda^*,\delta^*)} \left\{ -\Delta c(\lambda) + \frac{U^p + o(\Delta)}{1+r\Delta} + \frac{1}{1+r\Delta} \left[\lambda\Delta \int_{\delta^*}^{\bar{\delta}} (V^p - U^p)\, dG(\delta) \right] \right\}. \tag{2'}$$

The first-order conditions (F.O.C.) of maximization with respect to λ^* and δ^* are respectively :

[21] If physicians were capacity constrained, then even if a patient, in his search, finds a physician whose type is relatively high, there is still an uncertainty that this physician will provide care to him. I have examined the model with this new feature by assuming that there is an exogenous probability that if the patient finds a physician of a relatively high type, he will receive care from him. But the main intuition and results of the model remain almost unchanged. However, patients who are willing to switch have to search longer and the switching rate in every period will be lower than in the no capacity constraint setting.

$$-\Delta c'(\lambda^*) + \frac{1}{1+r\Delta}\Delta \int_{\delta^*}^{\bar{\delta}} \left[V^p\left(x, h(\theta,\delta)\right) - U^p(x, h(\theta,\delta^c))\right] dG(\delta) = 0 \tag{3}$$

and

$$\frac{\partial \left[\int_{\delta^*}^{\bar{\delta}} \left(V^p - U^p\right) dG(\delta)\right]}{\partial \delta^*} = 0 \tag{4}$$

which can be reduced respectively to[22] :

$$c'(\lambda^*) = \int_{\delta^*}^{\bar{\delta}} \left[V^p\left(x, h(\theta,\delta)\right) - U^p(x, h(\theta,\delta^c))\right] dG(\delta) \tag{$3'$}$$

and

$$V^p\left[x, h(\theta,\delta^*)\right] - U^p[x, h(\theta,\delta^c)] = 0. \tag{$4'$}$$

Equations ($3'$) and ($4'$) implicitly describe the optimal critical physician's type and search intensity. The term on the right-hand side of ($3'$) reflects the expected gain attributable to finding an acceptable alternative (Mortensen (1986)) (i.e., the expected gain when a patient finds a physician whose type is at least equal to the type of his current physician). Thus equation ($3'$) provides that the optimal search intensity is such that the marginal cost of search is equal to the expected gain generated by the optimal search strategy.

Equation ($4'$) states that in equilibrium, the critical physician type is such that the expected utility from search is equal to his present level of utility. Furthermore, since the

---

[22] Equation ($3'$) follows from the fact that if the time interval Δ is very small then $lim\left(\frac{1}{1+r\Delta}\right) = 1$. Equation ($4'$) follows from Leibniz's rule. In fact, it states that if $J(a,b) = \int_a^b F(t,x)dt$ where t is the variable of integration and x is a variable which is neither the variable of integration nor a limit of integration, then $\frac{\partial J}{\partial a} = -F(t,x)\Big|_{t=a} = -F(a,x)$ i.e. the derivative of a definite integral with respect to its lower limit of integration is the negative of the integrand evaluated at that point.

Hence, $\frac{\partial\left[\int_{\delta^*}^{\bar{\delta}}(V^p - U^p)dG(\tilde{\delta})\right]}{\partial \delta^*} = \frac{\partial\left[\int_{\delta^*}^{\bar{\delta}}(V^p - U^p)g(\tilde{\delta})d\tilde{\delta}\right]}{\partial \delta^*} = -\left\{\left[V^p(x,h(\theta,\delta)) - U^p(x,h(\theta,\delta^c))\right]g(\delta)\right\}\Big|_{\tilde{\delta}=\delta^*}$
$= -\left\{\left[V^p(x,h(\theta,\delta^*)) - U^p(x,h(\theta,\delta^c))\right]g(\delta^*)\right\}.$

Since $\frac{\partial V^p}{\partial \delta^*} = 0$ is equivalent to $\frac{\partial\left[\int_{\delta^*}^{\bar{\delta}}(V^p - U^p)dG(\tilde{\delta})\right]}{\partial \delta^*} = 0$, and using the fact that $g(\delta^*) > 0$, the equation (4') follows.

patient perceives his current physician's true type δ^c as δ^{EV}, then equation (4′) also states that at the optimum, the critical physician type is equal to the patient's perception of the type of his current physician.

In order to find a more tractable relationship between λ^* and δ^*, I integrate (3′) by parts, which yields :

$$c'(\lambda^*) = \left[V(x, h(\theta, \bar{\delta})) - U^p(x, h(\theta, \delta^c))\right] G(\bar{\delta}) - \left[V(x, h(\theta, \delta^*)) - U^p\right] G(\delta^*) - \int_{\delta^*}^{\bar{\delta}} G(\delta) V' d\delta.$$

Using (4′) and the fact that $G(\bar{\delta}) = 1$, we get :

$$c'(\lambda^*) = \left[V(x, h(\theta, \bar{\delta})) - U^p(x, h(\theta, \delta^c))\right] - \int_{\delta^*}^{\bar{\delta}} G(\delta) V' d\delta. \tag{5}$$

The term in the square brackets on the right-hand side of (5) is the difference between the levels of utility procured by the physician whose treatment level is the highest ($\delta = \bar{\delta}$) and the current one ($\delta = \delta^c$). The last term is the marginal gain from search when the patient changes physician weighted by the corresponding probabilities. In other words, equation (5) states that : at equilibrium, the optimal search intensity and the critical physician type are such that the marginal cost of search is the difference between the gain from searching for a physician with the highest type and the expected marginal gain attributable to finding a physician whose type is greater than δ^*.

Totally differentiating both sides of (3′) and using the earlier application of Leibniz's rule leads to :

$c''(\lambda^*) d\lambda^* = -\left\{\left[V^p(x, h(\theta, \delta^*)) - U^p(x, h(\theta, \delta^c))\right] g(\delta^*)\right\} d\delta^*$ and

$c''(\lambda^*) \frac{d\lambda^*}{d\delta^{EV}} = -\left\{\left[V^p(x, h(\theta, \delta^*)) - U^p(x, h(\theta, \delta^c))\right] g(\delta^*)\right\}.$[23]

Every patient who switches believes that his critical physician's type δ^* is such that $\delta^* \geq \delta^c$ or equivalently $V^p(x, h(\theta, \delta^*)) - U^p(x, h(\theta, \delta^c)) \geq 0$. Since $c'' \geq 0$ and $g(\delta^*) > 0$,

[23] Because, at the equilibrium the patient sets his critical physician's type δ^* equal to his evaluation δ^{EV} of his current physician's type then $d\delta^* = d\delta^{EV}$.

it follows that : $\frac{d\lambda^*}{d\delta^{EV}} \leq 0$ i.e., the patient's optimal search intensity is decreasing in his perception of the current physician's type. Consequently, if the patient has a physician whose treatment level is perceived as relatively high (low), then he has little (a lot of) incentive to search actively. Thus, depending on their perceptions of their physicians' type, patients with the same health condition and the same physician may search differently because of their different level of ignorance (i.e., different level of μ).

The analysis of the optimal search strategies from equation (4′) reveals that, a patient will leave his physician if and only if the expected value from search and the current utility are equal. This means that the critical physician type which will induce the patient to switch has the following characteristics : leaves if $\delta > \delta^*$ and stays if $\delta \leq \delta^*$.

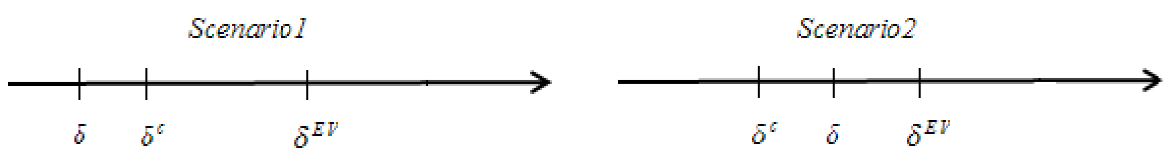
Proposition 1

Under perfect information about the outside physician's type, a patient will make an appropriate decision (i.e., the patient will stay with a physician who provides appropriate care and leave a physician who provides inappropriate care) if and only if he perfectly estimates his illness severity ; otherwise the appropriateness of the decision depends on the extent of the difference between the patient's evaluation of his current physician's type and the outside physician's type.

Proposition 1 states that if the patient has perfect information about the alternative physician, then his decision depends only on his level of ignorance about his illness severity. As a consequence, if the patient perfectly estimates his illness severity then his decision is appropriate. However, if the patient's estimation of his illness severity is imperfect then patients may make type I and II errors ; that is, some patients may stay with their physicians even if they do not receive the appropriate care while others may leave good physicians for poorer ones. In order to illustrate such errors, take the case of an under-estimator of his illness severity ($\mu < 0$) who does not switch and an over-estimator ($\mu > 0$) who does switch.

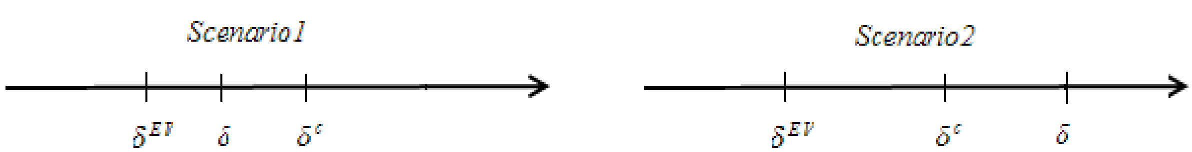
In the former case, if the type of the outside physician δ is close to the patient's evaluation of his current physician δ^{EV}, then the outside physician's type may be greater than the current physician's true type δ^c (see Scenario 2 in **Figure 1.2-a** below). As a consequence, the patient stays with a physician whose type is relatively low.

Figure 1.2–a : $\mu < 0$, Patients who Do Not Switch.



In the case of an over-estimator of the illness severity ($\mu > 0$) who switches, if the type of the outside physician δ is not far greater than the patient's evaluation of his current physician δ^{EV}, then the outside physician's type may be lower than the current physician's true type δ^c (see Scenario 1 in **Figure 1.2-b** below). As a consequence, the patient leaves a physician whose type is relatively high.

Figure 1.2–b : $\mu > 0$, Patients Who Switch.



In the next section, I relax the simplifying yet questionable assumption that the patient perfectly observes the outside physician's type. In fact, as noted previously, in practice patients can evaluate their physicians' type only after health-care services are consumed; consequently, it is unrealistic to assume that patients have more information about the outside physician than the current one. I will therefore assume that when the patient finds a new physician, he also collects informations about the type of this physician. The information collection process is given in more detail below.

3.2.2 The Case Where the Outside Physician's Type Is Imperfectly Observable

In this section, I assume that the patient cannot perfectly observe the alternative physician's type but can nonetheless collect information regarding it through a social network (or a similar system).[24] More specifically, after finding a potential alternative physician, a patient contacts a subset of the sample of people in his social network and asks for details about that specific physician. The information obtained represents a signal of the type of this particular physician, and is denoted by $\tilde{\delta}$.[25] It is assumed that all patients in this specific subset seek care from that particular physician or have already used his services. This assumption, however, does not mean that people in the social network are homogeneous. In fact, patients with different kinds of illnesses may receive services from a same physician and, independently of their illness, can give their overall appreciation about that particular physician. Also, even if it were the case that people in the social network have the same illness, their overall appreciation would depend on their level of ignorance (I allow such heterogeneity in the social network). I assume that $\tilde{\delta} = \delta + \varepsilon$ where ε is i.i.d with mean zero. More specifically, the signal obtained by the patient depends on the type of the people contacted in the social network. Hence if $\varepsilon > 0$, then the patient has collected the

[24] Social network is described in the literature as individuals or groups with whom a particular individual is in contact.

[25] At this stage, it is important to note that once a physician is contacted, the information obtained from the social network is assumed to be costless. In other words, I assume that the information collection process is exogenously given. One may however endogenize it by assuming that the signal obtained depends on the patient's search intensity through his social network. This limitation is left for future work.

information from a sample of patients who on average over-evaluate their physician's type while if $\varepsilon < 0$, the information is collected from patients who on average under-evaluate their physician's type. So, unlike the previous case where the patient is perfectly aware of the type of the outside physician, the patient now leaves if $\tilde{\delta} > \delta^*$ and stays if $\tilde{\delta} \leq \delta^*$. That is, the switching decision is now based on the signal of the alternative physician's true type. Although the patient knows that the signal may be biased, I assume that it is always better for him to follow his network's recommendation. To maintain consistency throughout I continue to assume that individuals are perfect, under or over-estimators of their illness severity. As a result, depending on the patients' type, three possible cases may arise. In what follows, each case is analyzed seperately.

Case 1 : Patients with Type $\mu = 0$

If $\theta^p = \theta$, then after health care consumption, the patient can attribute the entire difference between his expected treatment and his realized treatment to the physician and thus can perfectly infer his current physician type. In this case, patients' switching decision depends only on the type of the information collected about the outside physician.

Proposition 2

When the patient is a perfect estimator of his illness severity, he makes the appropriate decision if (i) he collects information from people who are on average over-evaluators of the outside physician's type and he does not switch, (ii) he collects information from under-evaluators of the outside physician's type and he switches ; otherwise the appropriateness of the decision depends on the extent of the patient's social network's mis-evaluation.

Proposition 2 states that a patient who is perfectly aware about his illness severity may nevertheless make an inappropriate decision, i.e., leave a physician whose type is relatively high for one whose type is relatively low, or stay with a physician whose type is relatively low. These errors result from the fact that the patient's social network evaluation of the alternative physician is imperfect.[26] If it were the case that the evaluation was perfect, then the search process would help the patient to make an appropriate decision.

I move now to a more interesting case where patients do not perfectly observe their true health condition. In this setting, the switching decision will depend not only on the type of the information givers but also on the type of the patient. That is, the switching decision will depend on both δ^{EV} and $\tilde{\delta}$.

Case 2 *:* Patients with Type $\mu < 0$

The patient under-estimates the illness severity and consequently he over-evaluates his physician's type. In other words, $\delta^{EV} > \delta^{c}$ and the optimal critical physician type will also be over-evaluated, leading patients to set their reservation type greater than their current physician's true type. This result follows from the fact that I assumed that patients evaluate their physicians' types by comparing their expected treatment to their realized treatment.[27] As a result, the patient's switching decision depends both on the extent of his over-evaluation of the physician's type and on the type of the information collected about the outside physician.

Proposition 3

When the patient is an under-estimator of his illness severity and for a given level of his evaluation (of his current physician), he makes the appropriate decision if and only if he collects information from people who are on average under-evaluators of the outside physician's type and he switches ; otherwise the appropriateness of his decision depends on the difference between the patient's evaluation and his social network's perception of the alternative physician, or on the extent of the patient's social network's mis-evaluation or

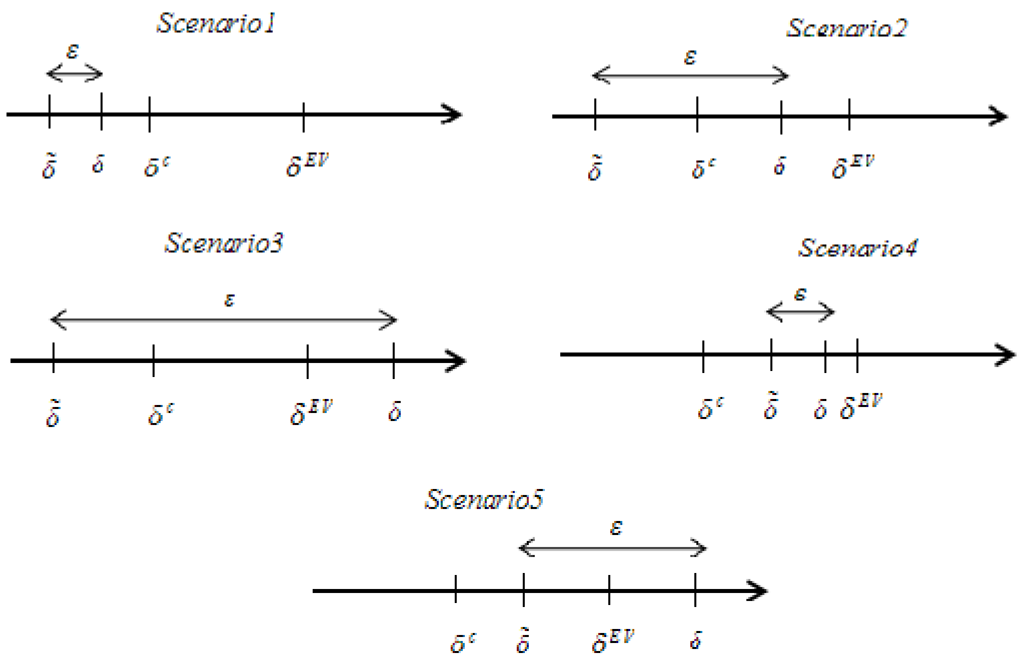
[26] In a multi-period setting, the patient may learn about the physician's type and as a consequence he is more likely to make an appropriate decision.

[27] If the evaluation is based on the comparison of the expected post-treatment health and realized post-treatment health, then a patient who under-estimates his illness severity may also under-evaluate his physician's type. This limitation is left for future work.

on both.

Proposition 3 states that search may be a good tool to find a better physician only if the illness severity is under-estimated and the signal about the alternative physician's type is obtained from under-evaluators (on average), and the patient switches. The main problem here is that the patient now faces two types of uncertainty : the ignorance about his health condition and the mis-evaluation by the social network. The combination of these uncertainties renders the patient's decision more complicated, and as such, renders him more likely to take inappropriate actions. To illustrate this complication, consider for example the case of an under-estimator of his illness severity ($\mu < 0$) who collects information from other under-evaluators of the alternative physician' type ($\varepsilon < 0$) and does not switch. He faces five scenarios (see **Figure 1.2-c** below) and his decision is appropriate only when scenario 1 arises.

Figure 1.2–c : $\mu < 0$ and $\varepsilon < 0$, Patients Who Do Not Switch.



Case 3 : Patients with Type $\mu > 0$

The patient over-estimates the illness severity and consequently under-evaluates his physician's type, i.e., $\delta^{EV} < \delta^c$. The optimal critical physician type will also be under-evaluated, that is, patients will set their reservation type lower than their current physician's true type. As in the two previous cases, I integrate the social network's information.

Proposition 4

When the patient is an over-estimator of his illness severity and for a given level of his evaluation (of his current physician), he makes the appropriate decision if and only if he *collects information from people who are on average over-evaluators of the alternative physician's type and he does not switch; otherwise the appropriateness of his decision depends on the difference between the patient's evaluation and his social network's perception of the alternative physician, or on the extent of the patient's social network's mis-evaluation or on both.*

As in the case of under-estimators (of their illness severity) and for similar reasons, patients face many scenarios and are more likely to take inappropriate actions.

To sum up, the patient's switching decision depends on his evaluation of the type of the current physician and on the signal of the alternative physician's true type. Depending on the extent of these elements, patients may make inappropriate decisions. In other words, some patients may stay with their physicians even if they do not receive the appropriate care while others may leave good physicians for poorer ones. That is, patients make type I and II errors.

In the next section, I endogenize the physicians' treatment decision as a function of the search behaviour of patients.

4 The Competition Between Physicians

In this section, I focus on physicians' responses given patients' search behaviour. I consider a two-period model in which physicians can costlessly diagnose their patient's illness severity. However, the physician bears a cost $C(\delta)$ per patient when he chooses a proportion δ of the appropriate treatment, where C is assumed to be an increasing and convex function. The proportion chosen is assumed to be in the support $\left[\delta^{\min}, \bar{\delta}\right]$, where $\delta^{\min}$ and $\bar{\delta}$ correspond respectively to the minimum and the maximum treatment level a physician can choose.[28] I assume for simplicity that, for a given illness severity θ, C is the only per-patient

[28] One may set the minimum level of δ at 0 and interpret it as if the physician refused to treat the patient. But this would be unrealistic in our context.

per-period cost borne by the physician. The physician's first period net income per-patient is thus given by $P - C(\delta)$, where P is the capitation fee as defined previously.

The physician is assumed to be a pure income maximizer (i.e., I rule out the possibility for elements such as altruism or ethical consideration or heterogeneity in the underlying talent of physicians). To ensure physician participation, the optimal payment design must give a minimum non-negative profit to the physician. Physicians are assumed to maximize their total (discounted) expected utility. Finally I assume that the physician's future revenue depends on the likelihood that the patient stays with him. This likelihood (which is detailed below) depends on the patient's outside option and on the type of people contacted in the social network. I continue to assume that patients (after their first-period health care consumption), evaluate their physician's type and subsquently decide on whether to stay with their current physician or seek care from a competing physician in the second period. If the patient decides to search, he follows the optimal search strategy derived in section 3.

4.1 The Physician's Behaviour

I showed in section 3 that a patient will wish to switch if his evaluation δ^{EV} of his current physician's type is less than a critical physician type δ^*. After searching, the patient will effectively switch if, in his search process, he receives a signal $\tilde{\delta}$ of an alternative physician which is greater than δ^*. Knowing this, the physician chooses δ to maximize his total expected utility W.

$$W = P - C(\delta) + [P - C(\delta)]\,\omega(\delta) \tag{6}$$

which can be rewritten as :

$$W = [P - C(\delta)]\,(1 + \omega(\delta)), \tag{6'}$$

where $\omega(\delta)$ is the probability that the patient will stay with the physician in the second period, which is assumed to be increasing in δ. From (6), one can see that a higher proportion of treatment reduces physicians' current profit but increases the likelihood that the patient stays with him. Without loss of generality, I assume that the discount factor is equal to one. It follows from (6′) that the expected utility for a physician who loses his patient in the second period is simply $W_0 = P - C(\delta)$ (i.e. $\omega(\delta) = 0$). More specifically $\omega(\delta)$ is defined as :

$$\omega(\delta) = \pi \left[\Pr(\delta^{EV} > \delta^*) + \Pr(\delta^{EV} < \delta^*) \left((1 - \lambda\Delta) + \lambda\Delta \Pr(\tilde{\delta} \leq \delta^{EV}) \right) \right] \tag{7}$$

which states that with probability π, the patient becomes ill in the second period and : (i) will not search for a competing physician if his evaluation of the current physician's type is greater than his critical physician's type (i.e., if $\delta^{EV} > \delta^*$) ; (ii) if $\delta^{EV} \leq \delta^*$, then the patient searches and with probability $(1 - \lambda\Delta)$ he does not find any alternative physician and therefore stays with his current physician ; (iii) finally, with probability $\lambda\Delta$ he finds a physician but his social network's average perception of that physician is lower than the patient's evaluation of his current physician's type and consequently, the patient stays with his current physician.

Denote $W_0^{\min} = P - C(\delta^{\min})$ as the physician's first period per-patient utility evaluated at $\delta^{\min}$, where $\delta^{\min}$ is the minimum proportion of the appropriate treatment a physician can choose. In order to rule out trivial solutions, I make the two following assumptions : (i) the capitation fee is set such that physician's utility is always non-negative under the appropriate treatment ; and (ii) the physician's first period per-patient utility evaluated at $\delta^{\min}$ is lower than the total expected utility evaluated at $\delta = 1$. The assumption (i) simply identifies the physician's participation constraint while (ii) ensures that the physician's optimal strategy is not to set the treatment level at its minimum.

4.2 Equilibrium Analysis

I solve the model by first assuming that the physician perfectly observes the patients' beliefs about their health condition (before the treatment occurs). I then relax this assumption later on.

4.2.1 The Case Where Patients' Type Is Observable

As I assume that the physician perfectly observes the patient's type, he would like to maintain the patient into the second period if, conditional on the patient's expectation about the value of δ, $W_0 \geq 0$ and $W > W_0^{\min}$. Under these conditions, the physician's total expected utility increases with the continuation of the relationship. Thus, his optimal choice is therefore to choose the treatment level which induces a greater level of satisfaction for the patient and consequently does not induce the patient to switch.[29] This strategy however would be optimal if and only if treatments were costless. The fact that the treatment is costly for the physician limits what he can or may be willing to do.

Given that the appropriate treatment corresponds to $\delta = 1$, a patient who under-estimates (over-estimates) his illness severity, expects to receive less (more) than 1. Since the treatment cost $C(\delta)$ is increasing in the proportion of treatment chosen, the physician has no incentive to choose more treatment than what the patient expects to receive from him. Moreover, choosing more than the appropriate level may make physicians' first-period per-patient profit negative.

Recall the assumption that the capitation fee is set such that, even if the physician provides the appropriate treatment (i.e., $\delta = 1$), he must still earn a non-negative profit. The physician's problem is to choose, for a given level of illness severity, the proportion of treatment which maximizes ($6'$) knowing that the probability that the patient will stay

[29] However, in the presence of a waiting list, the physician may decide not to maintain the patient into the second period.

with him into the second period is given by (7).

Proposition 5

When the patient's type is observable to the physician, then :

(i) the physician chooses a proportion of treatment which is lower than both the critical and the appropriate proportions, but may keep the patient into the second period if the patient is an under-estimator of his illness severity,

(ii) the physician may choose a proportion of treatment which is greater than the appropriate proportion but smaller than the critical one if the patient is an over-estimator of his illness severity.

Proposition 5 states that the physician will not treat the patient who under-estimates his illness severity appropriately and that such patient may not switch to another physician. This physician's behaviour is driven by two factors : (i) the patient who under-estimates his illness severity expects to receive less than the appropriate treatment ; and (ii) the probability that the patient leaves given his search behaviour is less than one. Proposition 5 also states that patients who over-estimate their illness severity may receive more than the appropriate treatment. This is because a patient who over-estimates his illness severity expects more than the appropriate treatment, and depending on the magnitude of the patient's ignorance about the illness severity, the physician may respond by providing more than the appropriate treatment. Thus, the model predicts a form of defensive medicine based on the patient's threat of seeking care elsewhere like in ALR (2009).

4.2.2 The Case Where Patients' Type Is Unobservable

In this section, the model is modified by allowing physicians to ignore the type of the patient they face. Since physicians are not able to perfectly observe their patient's type, I nonetheless allow them to form beliefs about their patient's type. That is, I assume that physicians know the distribution of patient-types (i.e., to what extent they over or under estimate their illness severity. I examine how this uncertainty about the patient's type affects the above results.

Proposition 6

When the patient's type is unobservable to the physician, he chooses a proportion of treatment which is lower than the appropriate proportion.

Proposition 6 states that if the patient is an under-estimator of his illness severity, he will receive more than what he expects but less than the appropriate treatment, and will stay with the physician. The over-estimator will receive also less than the appropriate treatment and will judge it insufficient, and consequently will search and may leave. Hence, uncertainty about the patient's type forces the physician to over-treat some types of patients. Moreover, providing the appropriate treatment does not insure that the patient does not leave.

To sum up, the treatment decision of the physician depends on the patient's ignorance about his illness severity. Also, the physician, knowing that in their search process, patients do not switch automatically (due to the fact that they do not always meet an outside physician, or that the signal received about the outside physician's type is lower than the type of their current one), may choose a proportion of treatment which is smaller than the patient's critical one and can keep the patient into the second period. That is, the treatment decision is not based on a cost-benefit analysis but rather on both the patient's level of ignorance and his outside options. As a consequence, patients who under-estimate their illness severity will receive a proportion of treatment which is smaller than the appropriate treatment and those who over-estimate their illness severity may receive more than the appropriate treatment. So, in equilibrium, the risk of loosing a patient may induce physicians to provide more than the appropriate treatment. That is, as in ALR (2009), the model predicts a new form of defensive medicine which is not associated with the traditional argument of the risk of lawsuits for medical malpractice.

Given the above results, the model suggests that any mechanism which reduces the patient's level of ignorance about his illness severity may be welfare increasing. In fact, since the physician's treatment decision is based on the patient's level of ignorance, a well-informed patient may receive the appropriate treatment.

5 Concluding Remarks and Policy Implications

The question that motivated this paper is whether a patient's search for a competing physician may encourage efficiency in the health-care sector even in the presence of a prospective payment mechanism. To this end, I build a two-period model where patients, after health care consumption, evaluate their physicians' quality, and may decide (following a new health shock) to stay or to search and switch. The main contribution of the model resides in the formulation of patients' outside options.

First, I examine patients' switching and search behaviour taking physicians' treatment decision as given. I find that due to their ignorance about the true illness severity, some patients may stay with their physicians even if they do not receive the appropriate care while others may leave good physicians for poorer ones. That is, patients make both type I and II errors. Considering patients who switch, I find that the treatment they receive in the second period depends in most of the cases on either the extent of their ignorance or on that of their social network's mis-evaluation, or on both. More specifically, a patient's search for an outside physician may induce (in some circumstances) the patient to patronize a physician whose type is relatively lower than that of his first period's physician and, as a result, the patient will receive a relatively low quality of care.

Second, examining physicians' response given patients' search behaviour, I show, under some assumptions, that the treatment provided depends on the physician's information about the patient's type. Hence, at the equilibrium, the fear of loosing a patient in a competitive market does not systematically induce the physician to treat appropriately the latter. More specifically, if the patient's type is observable to the physician, a patient who under-estimates his illness severity will be under-treated by the physician while a patient who over-estimates his illness severity may be over-treated. If, however, the patient's type is unobservable to the physician, the physician will systematically under-treat the patient.

The model suggests several avenues which may encourage the efficient provision and consumption of care. First, patients should be well informed about both their current physician's quality and that of the outside physicians. This can be done by making information on physicians' performance available to patients by establishing a physician quality reporting system which is accessible to the patients. Second, since the physician's treatment decision is based on the patient's level of ignorance about his illness severity, a mechanism which reduces this ignorance may induce the physician to provide a level of treatment which is close to the appropriate level given the true illness severity. That is, patients should be more informed about their health conditions by providing them better information about a particular illness as well as the treatments available. It is important that these information are obtained through a system certified by an evaluating organization for accuracy. Finally, in addition to the above avenues, reducing search costs increases the patient's gain from searching for an alternative physician.

Although the context of this paper may seem more suitable to a system where there is a private insurance market and enough physicians (like a US-type system), the present work is also applicable in a market characterized by the presence of a public insurance and a shortage of physicians (like Canadian system) because the level of competition across physicians and the patient's switching decision are endogenized.

I have assumed a two-period model, so search as a form of precaution is ruled out. The model can be extended in a multi-period setting to include this feature. Also, in the model, I have assumed that once a potential physician is found, the information obtained about the signal of this particular physician is exogenously given. This information collection process can be endogenized by assuming that the signal obtained about the type of the alternative physician depends on the patient's search intensity through his social network. These limitations are left for future work.

6 References

[1] Allard Marie, Léger Pierre Thomas and Rochaix Lise (2009) 'Provider competition in a dynamic setting', *Journal of Economics and Management Strategy* 18(2), 457-486.

[2] Arrow Kenneth (1963) 'Uncertainty and welfare economics of medical care ', *American Economic Review* 53(5), 941-973.

[3] Burdett K. and Coles M. (1999) : 'Long-term partnership formation : marriage and employment', *Economic Journal* 109(406), 306-34.

[4] Choné Philippe and Ma Ching-to Albert (2011) 'Optimal health care contract under physician agency', *Annales d'Économie et de Statistique*, issue 101-102, 229-256.

[5] Danzon Patricia (2000) 'Liability for medical malpractice' A.J. Culyer and J.P.Newhouse, eds. *Handbook of Health Economic*s, Amsterdam : Elsevier North-Holland, 1339-1404.

[6] Dranove David (1988) 'Demand inducement and physician-patient relationship, *Economic Inquiry* 26(2), 281-298.

[7] Ellis P. Randall and McGuire Thomas (1993) 'Supply-side and demand-side cost sharing in health care', *Journal of Economic Perspectives* 7(4), 135-151.

[8] Gourash Nancy (1978) 'Help-seeking : a review of the literature', *American Journal of Community Psychology* 6(5), 413-423.

[9] Hoerger T. J. and Howard L. Z. (1995) 'Search behaviour and choice of physician in the market for prenatal care', *Medical Care* 33(4), 332-349.

[10] Jack William (2005) 'Purchasing health care services from providers with unknown altruism', *Journal of Health Economics* 24, 73-93.

[11] Kaiser Family Foundation/Agency for Health Care Research and Quality, (2000) 'National survey on Americans as health care consumers : an update on the role of quality information'.

[12] Léger Pierre Thomas (2000) 'Quality control mechanisms under capitation payment for medical services', *Canadian Journal of Economics* 33, 564-88.

[13] Ma Ching-to Albert and McGuire Thomas (1997) 'Optimal health insurance and provider payment', *American Economic Review* 87, 685-704.

[14] Mortensen Dale (1982) 'The matching process as a noncooperative game', in J.J. McCall ed, *Economics of Information and Uncertainty,* 233-258.

[15] Mortensen, D.T. (1986) 'Job search and labor market analysis', in : O. AsHenfelter and R. Layard (Eds) Handbook of Labor Economics, Amsterdam : North-Holland, 849-919.

[16] Rochaix Lise (1989) 'Information asymmetry and search in the market for physicians services', *Journal of Health Economics* 8, 53-84.

[17] Rossiter Louis, Langwell Kathryn, Wan Thomas and Rivnyak Margaret (1989) 'Patients satisfaction among elderly enrollees and disenrollees in medicare health maintenance organization : Results from the national medicare competition evaluation', *Journal of the American Medical Association* 262(1), 57-63.

[18] Schlesinger Mark, Druss Benjamin and Tracey Thomas (1999) 'No exit ? The effect of health status on dissatisfaction and disenrollement from health plans', *Health Services Research* 34(2), 547-576.

[19] Sobero E. Melony, Andrew W. Dick, Zwanziger Jack, Mukamel Dana and Weyl Nancy (2003) 'The Effect of Capitation on Switching Primary Care Physicians', *Health Services Research* 38(1), 191-209.

[20] Wolinsky Asher (1987) 'Matching, search and bargaining', *Journal of Economic Theory* 42(2), 311-333.

7 Appendix

7.1 Appendix A : Proof of Equation (2′)

$$V^p = \max_{(\lambda^*,\delta^*)} \left\{ -\Delta c(\lambda) + \frac{1}{1+r\Delta} \left[\begin{array}{c} \lambda\Delta \left(\int_{\delta^*}^{\bar{\delta}} V^p dG(\delta) + G(\delta^*) \times U^p \right) \\ + \\ U^p - \lambda\Delta \times U^p + o(\Delta) \end{array} \right] \right\}$$

$$= \max_{(\lambda^*,\delta^*)} \left\{ -\Delta c(\lambda) + \frac{U^p + o(\Delta)}{1+r\Delta} + \frac{1}{1+r\Delta} \left[\begin{array}{c} \lambda\Delta \left(\int_{\delta^*}^{\bar{\delta}} V^p dG(\delta) + G(\delta^*) \times U^p \right) \\ + \\ -\lambda\Delta \times U^p \end{array} \right] \right\}$$

$$= \max_{(\lambda^*,\delta^*)} \left\{ -\Delta c(\lambda) + \frac{U^p + o(\Delta}{1+r\Delta} + \frac{\lambda\Delta}{1+r\Delta} \left[\int_{\delta*}^{\bar{\delta}} V^p dG(\delta) + (G(\delta^*) - 1)U^p \right] \right\}$$

We know that $G(\delta^*) = \int_0^{\delta^*} dG(\delta) \implies 1 - G(\delta^*) = \int_{\delta^*}^{\bar{\delta}} dG(\delta)$, and $(G(\delta^*) - 1)U^p = -U^p(1 - G(\delta^*))$ hence$(G(\delta^*) - 1)U^p = -U^p \int_{\delta^*}^{\bar{\delta}} dG(\delta)$. Finally it should be noted that U^p (the patient's current level of utility) is fixed and is independent of the alternative physician's type, so $U^p \int_{\delta^*}^{\bar{\delta}} dG(\delta) = \int_{\delta^*}^{\bar{\delta}} U^p dG(\delta)$. By using this fact, it follows that :

$$V^p = \max_{(\lambda^*,\delta^*)} \left\{ -\Delta c(\lambda) + \frac{U^p + o(\Delta)}{1+r\Delta} + \frac{1}{1+r\Delta} \left[\lambda\Delta \int_{\delta^*}^{\bar{\delta}} (V^p - U^p)\, dG(\delta) \right] \right\} \qquad \text{Q.E.D.}$$

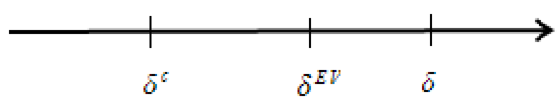
7.2 Appendix B : Proof of Propositions

Proof of Proposition 1

I show that the patient's switching decision depends on his evaluation of the type of the current physician δ^{EV} and on the type of the outside physician δ. Also, from the assumption about the patient's evaluation of his current physician's type, it follows that δ^{EV} is different from δ^c unless $\mu = 0$. As a consequence, a patient who perfectly estimates his illness severity (i.e. $\mu = 0$), compares δ^c to δ and decides whether to stay with his current physician or to switch. Hence, for patients who do not switch, δ^c is greater than δ and consequently $U^p(x, h(\theta, \delta^c) > V^p(x, h(\theta, \delta))$, while for those who switch, δ^c is smaller than δ and consequently $U^p(x, h(\theta, \delta^c) < V^p(x, h(\theta, \delta))$.

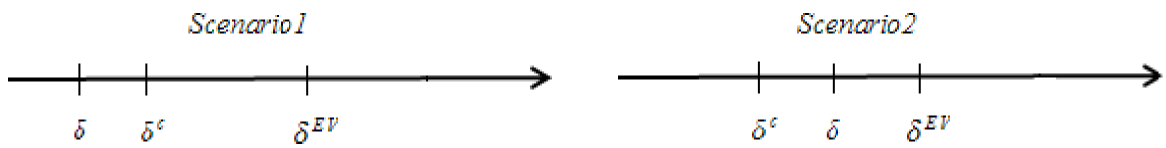
However, if the patient does not perfectly estimate his illness severity, then many scenarios may arise. More specifically, for a patient with $\mu < 0$, we know that $\delta^{EV} > \delta^c$. If this patient switches, then $\delta^{EV} < \delta$. Furthermore, since δ^{EV} is greater than δ^c, it follows that δ is always greater than δ^c (see **Figure 1.3-a**) and consequently $V^p(x, h(\theta, \delta)) > U^p(x, h(\theta, \delta^c)$.

Figure 1.3–a : $\mu < 0$, Patients Who Switch.



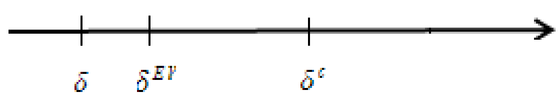
On the other hand, for patients who do not switch, and for a given level of evaluation, δ is smaller than δ^c if and only if δ is much smaller than δ^{EV} (see Scenario 1 in **Figure 1.3-b**)

Figure 1.3–b : $\mu < 0$, Patients Who Do Not Switch.



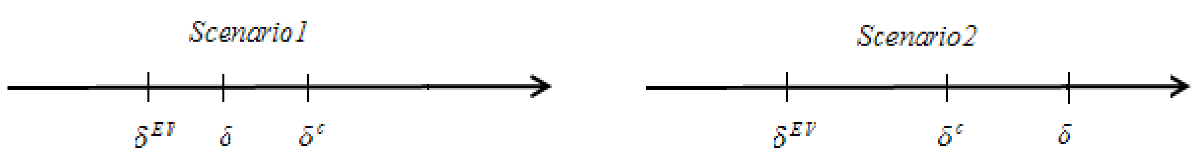
Now, if $\mu > 0$, we know that $\delta^{EV} < \delta^c$. For patients who do not switch, since $\delta^{EV} > \delta$ it follows that δ is always smaller than δ^c (see **Figure 1.4-a**).

Figure 1.4–a : $\mu > 0$, Patients Who Do Not Switch.



Finally, for those who switch, $\delta^{EV} < \delta$ and as a consequence δ is greater than δ^c if and only if δ is far greater than δ^{EV} (see Scenario 2 in **Figure 1.4-b**).

Figure 1.4–b : $\mu > 0$, Patients Who Switch.

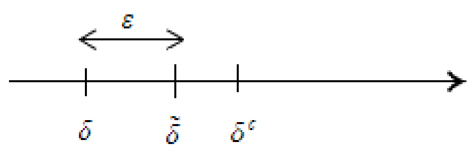


QED

Proof of Proposition 2

I have assumed that $\tilde{\delta} = \delta + \varepsilon$. So if the patient collects information from people who are on average over-evaluators of their physician's type (i.e., $\varepsilon > 0$) , then $\tilde{\delta}$ is greater than the true type δ of the alternative physician. Since patients base their switching decision on $\tilde{\delta}$, then for patients who do not switch, $\tilde{\delta}$ is smaller than δ^{EV} (which is equal to δ^c because $\mu = 0$). Hence, since δ is smaller than $\tilde{\delta}$, thus δ is always smaller than δ^c (see **Figure 1.5-a**) and consequently $V^p(x, h(\theta, \delta)) < U^p(x, h(\theta, \delta^c))$.

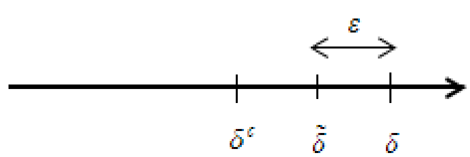
Figure 1.5–a : $\mu = 0$ and $\varepsilon > 0$, Patients Who Do Not Switch.



Thus patients stay with a relatively high type and the patient will receive a better care from the current physician than what he would receive if he switches.

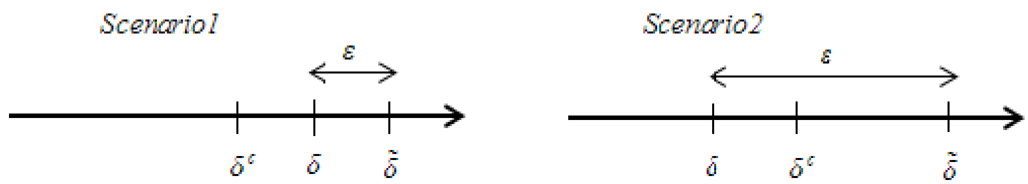
For those who switch, δ^c is smaller than $\tilde{\delta}$; and if the information is collected from people who under-evaluate their physician's type (i.e. $\tilde{\delta} \leq \delta$) then δ is always greater than δ^c whatever the value of ε (see **Figure 1.5-b**).

Figure 1.5–b : $\mu = 0$ and $\varepsilon < 0$, Patients Who Switch.



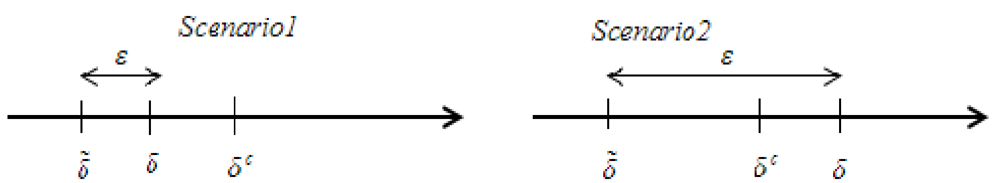
However, if $\varepsilon > 0$ and the patient switches, then his decision is appropriate if and only if ε is relatively small (see Scenario 1 in **Figure 1.6-a**).

Figure 1.6–a : $\mu = 0$ and $\varepsilon > 0$, Patients Who Switch.



Likewise, if $\varepsilon < 0$ but the patient stays with his current physician, then his decision is appropriate if and only if ε is relatively small (see Scenario 1 in **Figure 1.6-b**).

Figure 1.6–b : $\mu = 0$ and $\varepsilon < 0$, Patients Who Do Not Switch.

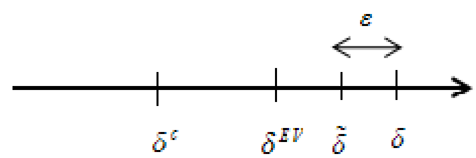


QED

Proof of Proposition 3

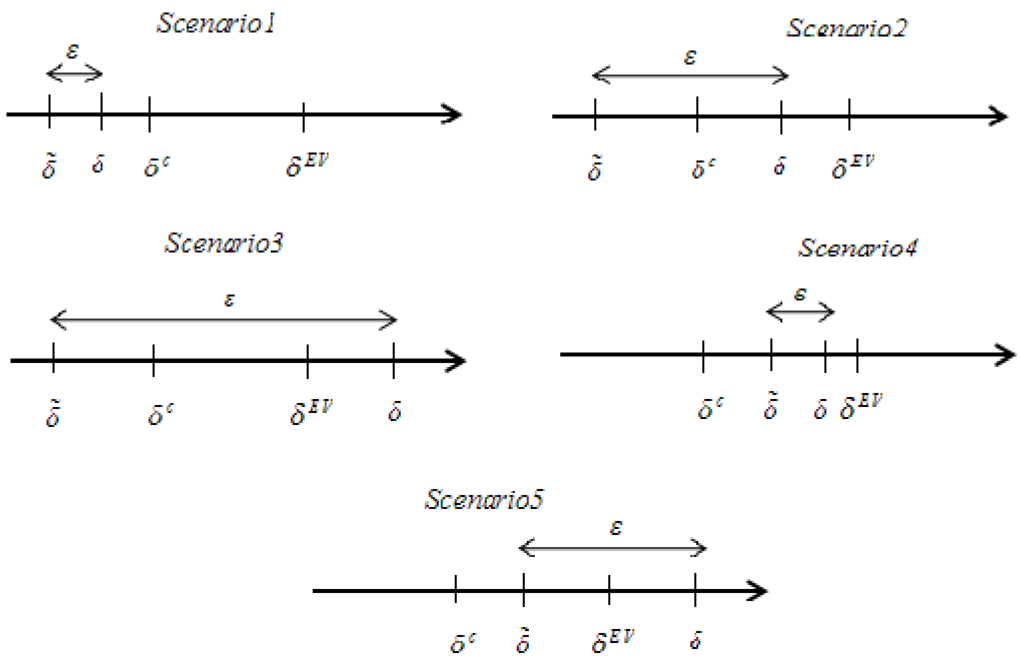
Recall that I examine here the case of a patient who is an under-estimator of his illness severity, i.e., $\mu < 0$. If the patient collects information from people who are on average under-evaluators of their physician's type (i.e., $\varepsilon < 0$), then $\tilde{\delta}$ is less than the true type δ of the alternative physician. We know that a patient switches if and only if $\tilde{\delta}$ is greater than δ^{EV}. Hence, since δ is greater than $\tilde{\delta}$, then for a patient who switches, δ is always greater than δ^c whatever the value of ε (see **Figure 1.7-a**) and consequently $V^p(x, h(\theta, \delta)) > U^p(x, h(\theta, \delta^c))$.

Figure 1.7–a : $\mu < 0$ and $\varepsilon < 0$, Patients Who Switch.



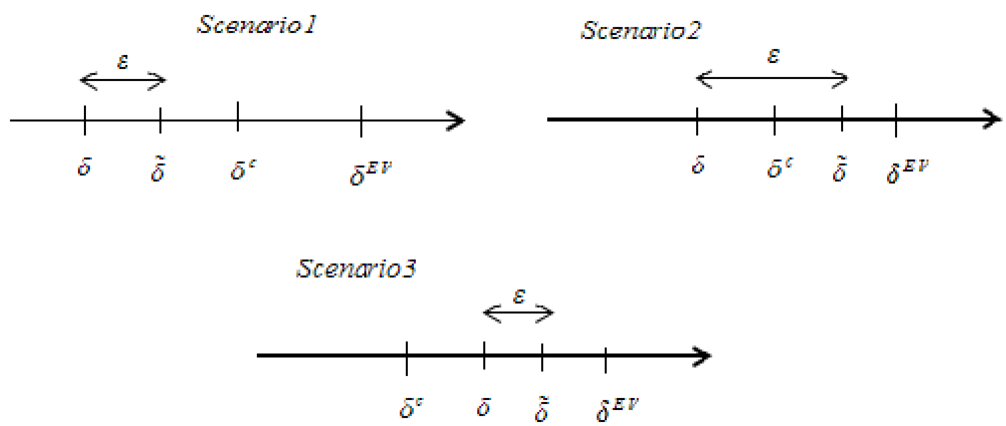
Patients who receive information from average under-evaluators of their physician's type but do not switch make an appropriate decision if and only if $\tilde{\delta}$ is much smaller than δ^{EV} and ε is very small (see Scenario 1 in **Figure 1.7-b**).

Figure 1.7–b : $\mu < 0$ and $\varepsilon < 0$, Patients Who Do Not Switch.



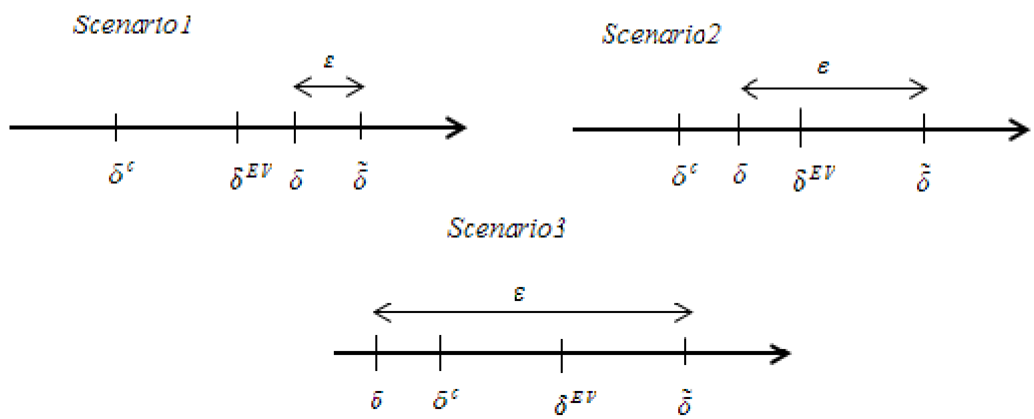
However, if $\varepsilon > 0$ and the patient does not switch, then his decision is appropriate if and only if $\tilde{\delta}$ is much smaller than δ^{EV}, or $\tilde{\delta}$ is close to δ^{EV} and ε is relatively high (see Scenarios 1 and 2 in **Figure 1.8-a**).

Figure 1.8–a : $\mu < 0$ and $\varepsilon > 0$, Patients Who Do Not Switch.



If however $\varepsilon > 0$ but the patient leaves his current physician, then his decision is appropriate if and only if ε is not too high (see Scenarios 1 and 2 in **Figure 1.8-b**).

Figure 1.8–b : $\mu < 0$ and $\varepsilon > 0$, Patients Who Switch.

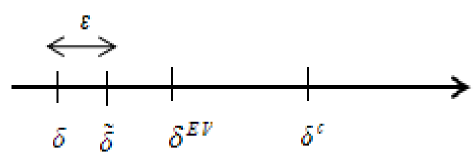


QED

Proof of Proposition 4

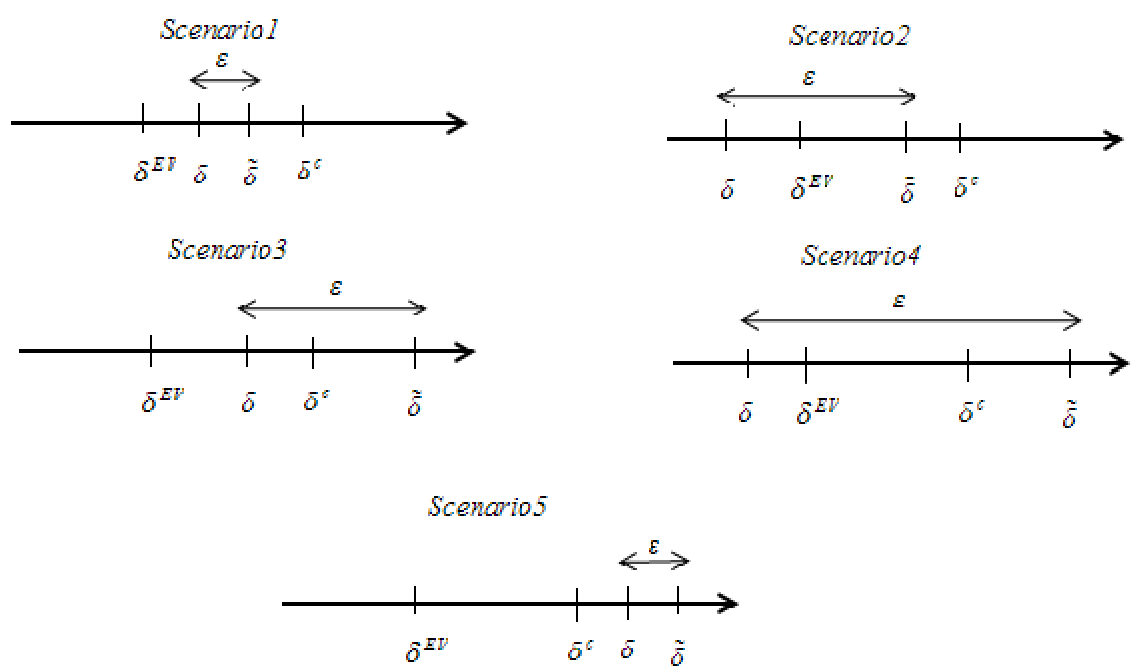
Recall that I examine the case of a patient who is an over-estimator of his illness severity, i.e., $\mu > 0$. If the patient collects information from people who are on average over-evaluators of their physician's type, then $\tilde{\delta}$ is greater than the true type δ of the alternative physician. We know that a patient switches if and only if $\tilde{\delta}$ is greater than δ^{EV}. Hence, for a patient who does not switch, since $\tilde{\delta}$ is greater than δ (because $\varepsilon > 0$) and δ^{EV} is smaller than δ^c (because $\mu > 0$) then δ is always smaller than δ^c (see **Figure 1.9-a**) and consequently $V^p(x, h(\theta, \delta)) < U^p(x, h(\theta, \delta^c))$.

Figure 1.9–a : $\mu > 0$ and $\varepsilon > 0$, Patients Who Do Not Switch.



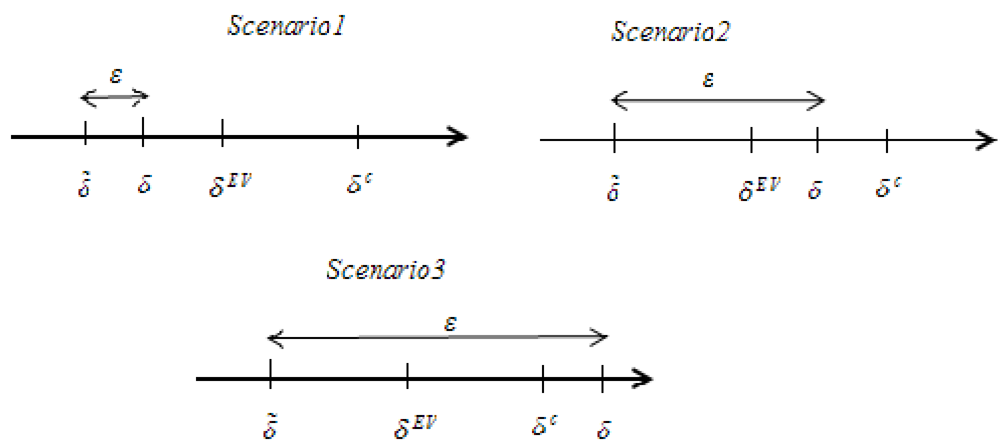
For a patient who receive information from average over-evaluators of their physician's type and switches, he makes an appropriate decision if and only if $\tilde{\delta}$ is much greater than δ^{EV} and ε is relatively small (see Scenario 5 in **Figure 1.9-b**).

Figure 1.9–b : $\mu > 0$ and $\varepsilon > 0$, Patients Who Switch.



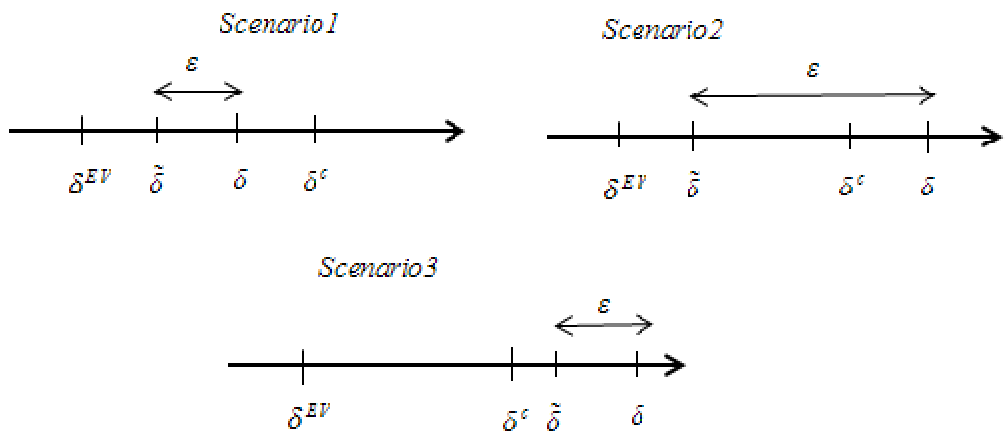
However, if $\varepsilon < 0$ and the patient does not switch, then his decision is appropriate if and only if ε is not too high (see Scenarios 1 and 2 in **Figure 1.10-a**).

Figure 1.10–a : $\mu > 0$ and $\varepsilon < 0$, Patients Who Do Not Switch.



If however, the patient leaves his current physician, then his decision is appropriate if and only if $\tilde{\delta}$ is not too close to δ^{EV} and ε is relatively high (see Scenarios 2 and 3 in **Figure 1.10-b**).

Figure 1.10–b : $\mu > 0$ and $\varepsilon < 0$, Patients Who Switch.



QED

Proof of Proposition 5

Since for an under-estimator of his illness severity, the critical proportion of treatment is lower than the appropriate one (i.e., $\delta^* < 1$), then $P - C(\delta^*) > P - C(1)$ and thus W_0 is always non-negative. The physician may provide to the patient who under-estimates his illness severity the appropriate treatment given the patient's estimate and keep him into the second period. But, since the physician knows that in their search process, patients do not switch automatically, he may choose a proportion of treatment which is smaller than the patient's critical one and may keep the patient into the second period even if the latter's outside options are unknown to the physician. Thus, for a given illness severity, the physician will choose a proportion of treatment such that δ is smaller than δ^*. And since $\delta^* < 1$, then at the equilibrium, δ will be smaller than both δ^* and 1, and must be between $\delta^{\min}$ and δ^*.

For an over-estimator of his illness severity, the equilibrium will be completely different. In fact, the physician knows that even if he provides to the patient the appropriate care, the patient will judge this to be insufficient and consequently will search and may switch. In this case, the physician's strategy will depend on the extent of the over-estimation of the illness severity. More specifically, even if choosing the patient's critical proportion of treatment δ^*, $P - C(\delta^*)$ is still positive, then the physician will choose the patient's critical proportion of treatment only if $P - C(\delta^{\min}) < [P - C(\delta^*)]\,(1 + \omega(\delta^*))$, otherwise, he will choose $\delta^{\min}$.[30] For the same reasons as in the case of an under-estimator, even if $P - C(\delta^*)$ is still positive, the fact that the switching probability when the patient searches is smaller than one gives some power to the physician and consequently he may choose a proportion of treatment which is smaller than what the patient expects (i.e., his critical proportion δ^*). That is, δ is less than δ^*. Since δ^* is greater than 1, hence, depending on the magnitude of μ, δ may be less than 1 or between 1 and δ^*. But if the patient's estimation is high enough such that $[P - C(\delta^*)]\,(1 + \omega(\delta^*)) \leq 0$, then the physician will always provide the minimum treatment and will loose the patient. QED

Proof of Proposition 6

Since the physician believes that, in expectation, the patient perfectly estimates his illness severity, he believes also that if he provides the appropriate treatment, he can keep him into the second period. But the physician knows that even if he does not choose the appropriate treatment, the patient may not switch due to the fact that the probability for a patient to find a physician whose perceived type is greater than hers is less than one. So the equilibrium proportion of treatment will be smaller than the appropriate one. QED.

[30] $P - C(\delta^{\min}) < [P - C(\delta^*)]\,(1 + \omega(\delta^*))$ means that the gain from choosing the minimum treatment and loosing the patient is lower than the gain from choosing the patient's desired level of treatment.

Essay 2
Physician Payment and Medical Malpractice Mechanisms

Abstract

In this paper, I build a model in which both physician payment and medical liability mechanisms may serve as a way to provide health services efficiently. The model predicts findings that are consistent with real world observations. First, it predicts that the fee incurred by the patient decreases the suing probability and second the patient's compensation is proportionate to the damage. I find that some physicians who treat appropriately their patients may be sued and thus may practice defensive medicine. I find that the risk aversion on patients' side lowers the claiming rate beyond efficient levels. I also find that the patient is always better off if the suing responsibilities are transferred to a law firm ; but the latter will only accept cases involving a seriously injury. The results suggest that an insurance policy for patients for malpractice suits may be welfare increasing.

1 Introduction

The patient-physician relationship is often characterized by asymmetric information. This characteristic may lead physicians to recommend or to provide the quantity and quality of care that not consistent with what a well informed patient would choose. To encourage efficiency, supply-side policies such as physician payment mechanisms have been put forth and they have received a lot of attention in the theoretical health economics literature. Another potential instrument is monitoring and medical malpractice mechanisms.[31] In fact, since the physician is responsible for both health cost containment and quality decisions (Blomqvist, (1991)), the utilization of two instruments may be more efficient. However, little work or policies have sought to combine these instruments which may be interdependent. In fact, there is an inter-dependency between the payment mechanism and the malpractice framework. That is, a particular malpractice rule or system may have different results if it is applied to different payment mechanisms. In light of this critique, the aim of this paper is to derive simultaneously optimal payment contracts for the physician and optimal compensation for the patient given that he is injured by the physician's action.

Each payment scheme is associated with its benefits (or advantages) but also its disavantages. The traditional fee-for-service (FFS) payment mechanism where (almost) all health-care costs reported by the physician are covered by the health insurance provider, gives strong incentives (in a static framework) to provide high-quality (or at least volume) care but no incentives to limit health-care costs, and health services may be provided beyond efficient levels. The prospective payment system (PPS) where the health insurer pays a fixed amount per patient enlisted into the physician's practice, may be an effective way to deal with excessive health-care costs. However, the PPS may give strong incentives to under-

[31] The monitoring mechanism is referring to verification made by an internal member while the malpractice mechanism is referring to that made by an external member.

provide health care and thus to provide low-quality care especially in static models. As a result, the literature on the design of optimal payment mechanisms for providers concludes that a combination of both FFS and PPS may be optimal. That is a mixed payment scheme may be efficient (e.g., Ellis and McGuire, (1986, 1990) ; Ma and McGuire, (1997) ; Selden, (1990)). These results, however, are sensitive to the models' assumptions. More specifically, Ellis and McGuire (1990) for example, assume that patients and physicians share the same level of information, that is, they do not model the traditional information problems between patients and physicians.

Although a few papers in the literature consider the role of two instruments, they do so in a framework where either liability mechanisms (Gal-Or, (1999) ; Zeiler, (2008)) or physician payment systems (Léger, (2000) ; Arlen and MacLeod, (2005)) are exogenously given.[32] In this paper, I also examine a model in which these two instruments may serve as a way to efficiently provide health services but in a more flexible payment mechanisms setting. More specifically, I examine how liability mechanisms influence the behaviour of the health insurer, the physician and the patient in a framework where physicians are paid through a mixed payment scheme. Unlike previous models, in the model proposed here, both physician payment and medical liability mechanisms are endogenized.

Initially, by assuming that all agents in the model are risk neutral, I find as in Léger (2000) and Zeiler (2008) that, at the equilibrium, the physician will never treat appropriately with certainty and as a consequence the patient also will never sue with certainty. However, in the present model, the flexibility of the payment mechanisms improves the patient's welfare in comparison with a capitation system. I find that the possibility of court errors

[32] In the theoretical literature on the optimal liability (see for example Becker, (1968) ; Shavell, (1980, 1991)), the main goal of liability is to provide incentives for an injurer to behave appropriately. Similarly, in the medical malpractice literature (see for example Danzon, (1985, 2000)), malpractice liability plays two roles : it gives incentives to physicians to provide appropriate care and compensates patients for injuries caused to them by physicians' inappropriate behaviour.

induces some physicians to practice defensive medicine. The model also predicts findings that are consistent with the real world observations. First, it predicts that the fees incurred by the patient to sue (in equilibrium) a physician are negatively related to the suing probability ; and second, the patient's compensation is proportionate to the damage. By relaxing the risk neutrality assumption on patients' side, I find that some patients who are treated inappropriately will not sue and as a consequence, the claiming rate will be low relative to efficient levels. That is, the results suggest that medical malpractice insurance from a patient's perspective might be welfare improving. I also find that a risk-averse patient will always prefer to transfer the suing responsibilities to a law firm ; but the latter will only accept cases involving a serious injury. In fact, given that the level of the patient's compensation is proportionate to the harm, thus the cases involving a seriously injury are more profitable.

The model is built on the approach taken by Zeiler (2008). The author, in a game-theoretical model, studies the effects of contract disclosure rules on the behaviour of health care markets' actors. In her model, the Managed Care Organization (MCO) chooses a physician payment mechanism and the physician, after observing the type of contract, decides whether to treat compliantly the patient or not. After health care consumption and upon the realization of a bad outcome, the patient forms beliefs about his physician's choice of treatment and decides whether to sue or not the latter for medical malpractice. If the patient sues, the court perfectly verifies the physician's action and sets the amount of damages if the physician is found liable. She finds under some assumptions that the regime in which MCO/physician contract terms are publicly observed is more efficient, i.e., the MCO can select a compensation scheme such that claiming rates will be lower and compliant treatment rates will be higher. In her model, the level of fine paid by the physician to compensate the patient is exogenous. The present work extends Zeiler's model by endogenizing the level of the compensation. Also, the model differs from Zeiler's by allowing the presence of court errors and the risk aversion on patients'side as in Léger, (2000).

The paper is also related to Gal-Or (1999). In a theoretical model where the health insurer is assumed to take into account both patients and physicians' welfare, and where the physician who is a pure profit maximizer chooses an intensity of the treatment and a health care cost control effort, she derives the optimal reimbursement and payer-provider sharing rules of malpractice liabilities. She shows that the optimal contract depends on the insurer's ability to control the provider's choice of the treatment. More specifically, if the insurer can dictate the intensity of the treatment, then the optimal contract is the prospective payment system and all malpractice liability costs are covered by the insurer. If the insurer cannot enforce the type of treatment, a mixed payment system may be optimal and the insurer shares both treatment and liability costs with the provider. The present model differs from Gal-Or's formulation in several aspects. First, Gal-Or assumes that the physician is found liable whenever his patient does not recover. In my model, I assume that the liability probability depends also on the patient's beliefs about whether he was treated appropriately or not (as in Zeiler 2008), and on the court's ability to verify the physician's actions. In other words, the probability of liability depends not only on the physician's behaviour, but also on other parameters. Second, in Gal-Or's paper, it is the health insurer and not the court or the regulator who plays the role of inducing efficient behaviour on the part of the provider and thus the fine is considered as exogenous by the insurer. In the model proposed here, I assume that the fine is endogeneous and is chosen by the regulator (social planner) in order to induce the physician to behave optimally.

The paper is also related to Léger (2000) which examines the physician's treatment decision and the patient's litigation decision in a model where physicians are paid by capitation.

The present work extends Léger's model by endogenizing the payment mechanism. Finally, the paper is related to Arlen and MacLeod (2005) who derive the optimal liability rules in a setting where both the physician and the health insurer's actions determine the care received by the patient. However, their model is silent on physician payment mechanisms.

The remainder of this paper is organized as follows. Section 2 presents the model. In section 3, I solve the model and present the results. In section 4, I allow for the possibility of court errors, i.e., the existence of type 1 and 2 errors. In section 5, I compare two different systems : the plaintiff is a risk-averse patient versus a risk-neutral law firm. Finally, the last section concludes.

2 The Model

In this section, I present a model in which many health care market actors interact. In the model, the health insurance provider collects premiums and pays physicians. I consider several forms of payment mechanisms. The physician's payment mechanism is either capitation, fee-for-service, or a mix of both. I assume that the health insurer signs contracts with the physician and the patient before the patient's illness severity is revealed. If an individual becomes ill, then he seeks care from the physician. The physician perfectly and costlessly observes the patient's illness severity and decides which type of treatment must be provided (so bad outcomes are not the result of ignorance but rather bad treatment). In the model, it is assumed that only the physician knows the true health condition of the patient. The other agents know only the distribution of the illness severity. Once health-care services are provided, the patient either recovers or not his previous health state. If he does not recover, the patient decides whether or not to sue the physician. If the patient decides to sue, then he pays fixed lawsuit costs and the physician's action is audited. The patient's true illness severity is (costly) revealed. If the court discovers that the treatment provided by the physician was inappropriate, then it will set endogenously the optimal level of fine that the physician must pay as compensation for the patient. Initially, I assume that the court can perfectly observe the patient's true illness severity. This assumption will be relaxed later on.

I describe in the following sections the timing of the model and the agents' preferences.

2.1 The Timing

The timing of the game (see Figure 2.1) is as follows :

Step 1

The health insurance provider offers contracts to the physician and the patient. These contracts specify the payment parameters (a fixed payment ρ and a cost-sharing rate γ) for the physician and the health insurance parameters (an insurance premium α and a co-payment rate β) for the insured individual.

Step 2

The patient becomes ill with either a low-severity illness θ_L which occurs with probability p, or a high-severity illness θ_H which occurs with probability $(1-p)$. As in Léger (2000), a patient with a high-severity illness requires a treatment t_H which costs $C(t_H) = C_H$ to the physician. I refer to this as the expensive treatment. While a patient with a low-severity illness requires a treatment t_L which costs $C(t_L) = C_L$ to the physician, which I refer to as the inexpensive treatment.

Step 3

The physician perfectly observes the patient's illness severity θ and decides on whether or not to treat appropriately the patient.[33] Once health care services are provided, the

[33] It is important to note that in the model proposed here, the treatment is assumed to be appropriate if a patient with low (high) severity illness receives the inexpensive (expensive) treatment. Otherwise, the treatment is inappropriate.

patient's post-treatment health is revealed and is either good or bad. If the patient is in a bad state, then he infers the probability (detailed below) that he was treated appropriately or not and decides either to sue or not the physician. So, it is assumed that malpractice here is always strategic and is not based on the physician's ignorance.

Step 4

Following the decision to sue the physician, the patient pays a fixed lawsuit costs (monetary costs to pay the lawyers) and his true illness severity is revealed. If the court discovers that the physician behaves inappropriately, then the court sets the optimal level of compensation.

2.2 The Patient

As noted above, the patient knows that he is sick but he does not know his illness severity. However, he knows the distribution of the severity : with probability p his illness is of low severity, and with $(1-p)$ the illness is of high severity. The patient, in the model proposed here, is assumed to be passive in the sense that he always follows his physician's treatment recommendation. After health care consumption, if the patient is in a good state then the game ends. If however the bad state occurs, then he may decide either to sue or not his physician. More specifically, given that the patient is in a bad state, if he observes that $t = t_L$ was chosen then he sues with probability f. If however $t = t_H$ is observed, then the patient sues with probability g. The patient's utility is assumed to depend on two elements : health state h and income available for consumption goods x, and is increasing and concave in both elements. The patient's utility function is thus given by :

$$U^P = U\left[h(\theta, t), x\right] \tag{8}$$

where $x = y - \alpha - \beta C(t)$; y is the individual's state-independent income.

2.3 The Physician

As noted above, the physician perfectly observes the patient's illness severity θ and decides which type of treatment must be provided. As in Léger (2000), once the patient's true illness severity is observed, the physician decides whether to provide the expensive treatment or the inexpensive one. More specifically, I assume that, given that the patient's illness severity is θ_L, the physician chooses t_L with probability u and t_H with probability $(1-u)$. If the severity is θ_H, I assume that the physician chooses t_L with probability v and t_H with probability $(1-v)$. In order to take into account the uncertainty between the effect of health care services' utilization and the health outcome, I assume that providing the appropriate treatment does not guaranty a full recovery of the patient's health. Also, even if the patient is inappropriately treated, he may recover his health. More specifically, I assume that, if $\theta = \theta_L$ and the physician chooses $t = t_L$ then the patient recovers with probability m; if however $t = t_H$ is chosen, then the recovery probability is n. Similarly, if $\theta = \theta_H$ and the physician chooses $t = t_L$ then the patient recovers with probability r; if however $t = t_H$ is chosen, then the recovery probability is s. The physician receives a fixed fee ρ which does not depend on the type of treatment provided and a share γ of the treatment cost reported to the insurer. The physician's utility function depends on whether or not he is sued. When he is not sued, the utility function is given by :

$$V^P = V\left[\rho - (1-\gamma)C_t\right] \tag{9}$$

where $\rho - (1-\gamma)C_t$ represents the financial gain : the physician receives ρ as a fixed payment, pays C_t for the treatment, and the third party reimburses γC_t (i.e., $\rho - C_t + \gamma C_t = \rho - (1-\gamma)C_t$).

However, if the physician is sued and the court discovers that the treatment provided was inappropriate, then the physician's utility is given by :

$$V^P = V\left[\rho - (1-\gamma)C_t - \phi_i\right] \tag{10}$$

where ϕ_i, $i = 1, 2$ denotes the fine paid by the physician to the patient given that the court finds that the physician behaves inappropriately. More specifically, ϕ_1 denotes the fine if the physician provides the inexpensive treatment given that the patient has a high-severity illness and ϕ_2 the fine when the expensive treatment is provided to a patient who has a low-severity illness. It is important to note that, if the physician is sued but the court discovers that the treatment provided was appropriate, then the physician's utility function is given by equation (9).[34]

2.4 The Third Party

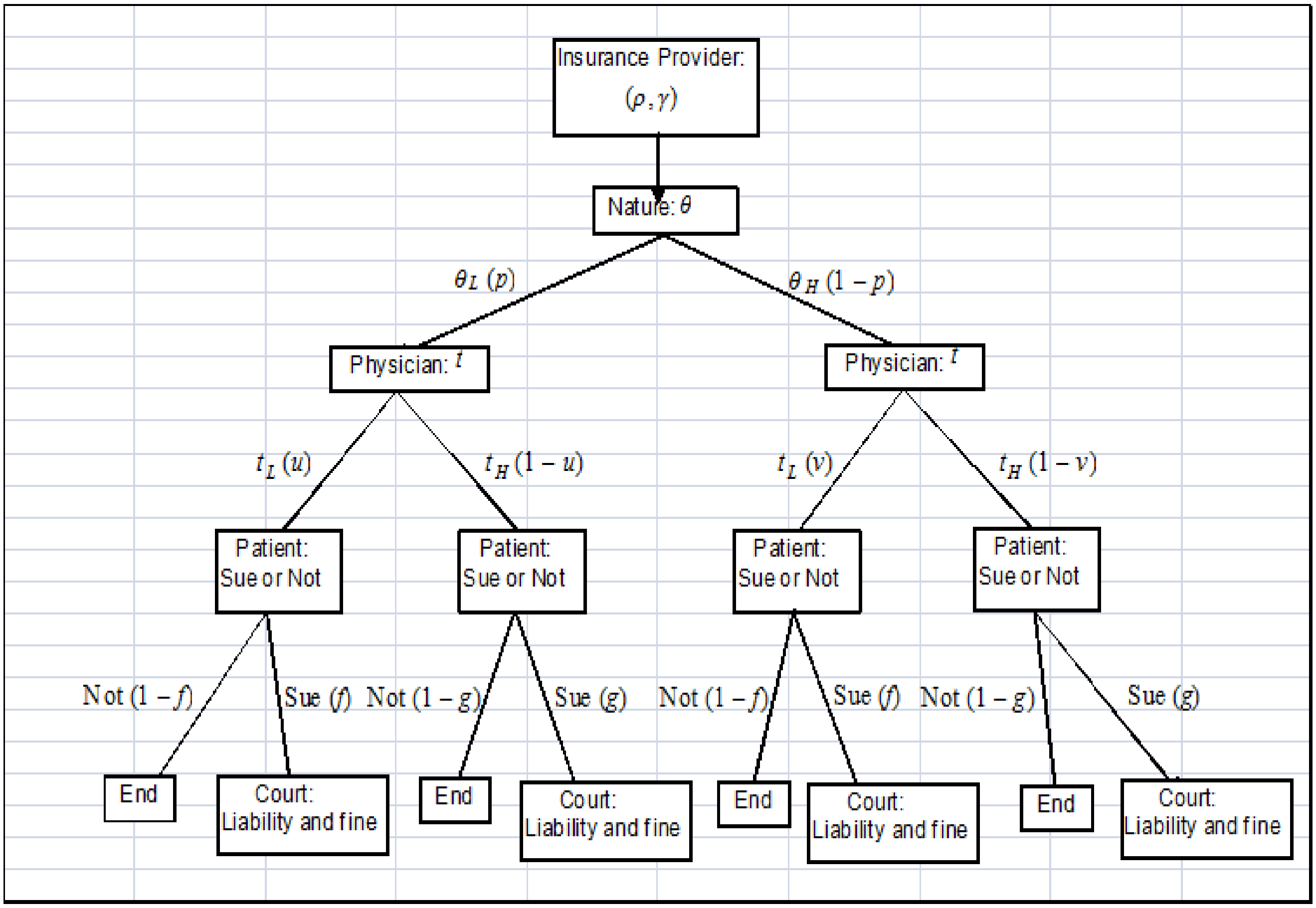
The third party is thought of as an insurance provider and is assumed to operate in a competitive market. The third party, which is assumed to be a profit maximizing firm, collects premiums and pays physicians by several forms of payment mechanisms. It chooses contract parameters to maximize its expected profits.

2.5 The Court

The court acts as a regulator of health care. It chooses ϕ_i to maximize social welfare. More specifically, the court maximizes the patient's expected utility subject to the constraints that the optimal liability mechanism must give a non-negative profit to the physician and the third party.

[34] I assume that, for the physician, there is no disutility associated with being sued. This assumption seems reasonnable because liability insurance premiums are not generally based upon a physician's record and are not increased for a history of complaints or on account of claims paid. However, physicians incur other financial and psychic costs (loss of time, loss of reputation) when they are sued for malpractice. Incorporating these costs into the model may result in more over-treatment. In fact, as suggested by Danzon (2000), the hidden costs of malpractice, such as psychic costs and loss of time defending malpractice allegations, are the primary drivers of defensive medicine.

Figure 2.1 : The Timing.



3 The Agents' Problem and Equilibrium Analysis

In this section, I formally write the maximization problem of the agents in the model and solve for the equilibrium behaviour of each agent using the Bayesian Nash equilibrium concept. I assume that all agents in the model are risk neutral. Later, I will relax the risk-neutrality assumption on the patient's side. I also make some simplifying assumptions : (i) the patient is fully insured ($\beta = 0$) and (ii) the physician cannot be fined by the court for an overtreatment $(\phi_2^* = 0)$. Given that $C_L < C_H$, the probability that the physician will treat a patient who has a low-severity illness with the expensive treatment is zero. As a consequence, the probability that the physician will appropriately treat a patient who has a low-severity illness is one, that is $u^* = 1$. Also, since the patient observes the treatment chosen by the physician, if he observes t_H he will never sue (i.e., $g^* = 0$). This is because the patient knows that $C_L < C_H$ and as a consequence the physician will never choose t_H for θ_L. As the above cases are ruled out, the physician and the patient strategies are reduced to choosing v and f, respectively. Through sections 3.1 to 3.5, I solve the model. The equilibrium parameters are found as follows : (i) I solve for the patient's, the physician's and the third party's optimal response given the level of the fine ; (ii) I resolve the court's problem given the behaviour of the patient, the physician and the third party ; (iii) I substitute the optimal fine into the optimal responses of the other agents. In section 3.6, I compare a system with a mixed payment mechanism with that of a pure capitation and a pure FFS.

3.1 The Physician's Behaviour

The physician who is assumed to be a pure income maximizer, chooses the cheating probability (v) to maximize his total expected utility. If the physician observes $\theta = \theta_H$ he may choose $t = t_L$ or $t = t_H$. Formally, conditional on observing $\theta = \theta_H$, the physician will choose $t = t_H$ if and only if :

$$V\left[\rho - (1-\gamma)C_H\right] > V\left[\rho - (1-\gamma)C_L - (1-r)f\phi_1\right] \tag{11}$$

where f denotes the probability that the patient will sue given that he was treated with t_L.

By assuming that the physician's utility is linear in his revenue, then :

$$f > \frac{(1-\gamma)(C_H - C_L)}{(1-r)\phi_1} = \bar{f} \tag{11'}$$

where $\bar{f}$ denotes the suing probability which makes the physician indifferent between treating appropriately or not given that the patient has a high-severity illness. Equation (11′)

states that if the probability that the patient will sue is higher than the suing probability which makes the physician indifferent between treating appropriately or not ($f > \bar{f}$), then the physician will always provide the appropriate treatment ($v = 0$), otherwise (i.e. if $f < \bar{f}$), he will never treat appropriately ($v = 1$).

3.2 The Patient's Behaviour

The patient's problem is to decide whether or not to sue. Before writing the maximization problem, we need to define the probability that a patient has a illness severity $\theta = (\theta_L, \theta_H)$ given that the treatment $t = (t_L, t_H)$ was provided. Let $P_{ij} = P(\theta = \theta_i / t = t_j)$ denote the probability that the patient has severity i given that the treatment j is observed, $i, j = L, H$. Using Bayes' rule, we have four cases :

$P_{LL} = \frac{pu}{pu+(1-p)v}$; $P_{HL} = \frac{(1-p)v}{pu+(1-p)v}$; $P_{HH} = \frac{(1-p)(1-v)}{(1-p)(1-v)+p(1-u)}$; $P_{LH} = \frac{p(1-u)}{(1-p)(1-v)+p(1-u)}$.[35]

Using the fact that at the equilibrium $u^* = 1$, these probabilities can be rewritten respectively as :

$P_{LL} = \frac{p}{p+(1-p)v}$; $P_{HL} = \frac{(1-p)v}{p+(1-p)v}$; $P_{HH} = 1$; $P_{LH} = 0$.

I assume that the patient's utility function is linear in its arguments. If the patient observes the inexpensive treatment (t_L), the patient will sue (given that he is in a bad

[35] The probability that $t = t_L$ is the sum of the probability that a patient with low severity illness is treated with the inexpensive treatment (pu) and the probability that a patient with high severity illness is treated with the inexpensive treatment ($(1-p)v$). Similary, the probability that $t = t_H$ is the sum of the probability that a patient with high severity illness is treated with the expensive treatment ($(1-p)(1-v)$) and the probability that a patient with low severity illness is treated with the expensive treatment ($p(1-u)$).

state) if and only if [36] :

$$v > \frac{p\,(1-m)\,k}{(1-p)\,(1-r)\,[\phi_1 - k]} = \bar{v} \tag{12}$$

where k represents the fee incurred by the patient to sue a physician and $\bar{v}$ denotes the cheating probability which makes the patient indifferent between suing or not given that he observed that the inexpensive treatment is provided. Equation (12) states that if the probability that the physician will provide the inappropriate treatment is higher than the threshold's probability which makes the patient indifferent between suing or not ($v > \bar{v}$), then the patient will always sue ($f = 1$), otherwise (i.e. if $v < \bar{v}$), he will never sue ($f = 0$).

3.3 The Patient and Physician Equilibrium Strategies

As shown by Léger (2000) and Zeiler (2008), at the equilibrium, patients and physicians use mixed strategies.[37] In fact, from the patient's behaviour, we state that if the probability that the physician will cheat is higher than the threshold probability which makes the patient indifferent between suing or not ($v > \bar{v}$), then the patient will always sue ($f = 1$). The physician, knowing that the patient will sue with probability 1, will choose $v \leq \bar{v}$. Hence $v > \bar{v}$ cannot be an equilibrium. If however, $v < \bar{v}$, then the patient will never sue, ($f = 0$). Knowing that the patient will never sue, the physician will choose $v \geq \bar{v}$. Thus $v < \bar{v}$ also cannot be an equilibrium. The physician's equilibrium strategy is $v^e = \bar{v}$.

Also, from the physician's behaviour, we know that if the probability that the patient

[36] Given t_L, if the patient sues, his expected utility is :

$$P_{HL}\,(1-r)\left[\underline{h}_H^L + y - \alpha - k + \phi_1\right] + P_{LL}\,(1-m)\left[\underline{h}_L^L + y - \alpha - k\right],$$

where $\underline{h}_i^L$ is the health state associated with the bad outcome given that a patient with illness severity i is treated with t_L.

If the patient does not sue, his expected utility is :

$$P_{HL}\,(1-r)\left[\underline{h}_H^L + y - \alpha\right] + P_{LL}\,(1-m)\left[\underline{h}_L^L + y - \alpha\right]$$

Substituting for P_{HL} and P_{LL}, equation (12) follows.

[37] Other equilibriums may exist. For details on them, see Léger, (2000).

will sue is higher than the suing threshold which makes the physician indifferent between treating appropriately or not ($f > \bar{f}$), then the physician will provide the appropriate treatment ($v = 0$). But, knowing that the physician will provide the appropriate treatment, the patient will choose $f \leq \bar{f}$. Thus $f > \bar{f}$ cannot be an equilibrium. If $f < \bar{f}$, then the physician will provide the inappropriate treatment ($v = 1$). The patient, knowing that the physician will never provide the appropriate treatment, will choose $f \geq \bar{f}$. Hence, $f < \bar{f}$ also cannot be an equilibrium. The patient's equilibrium strategy is $f^e = \bar{f}$.

Thus at the equilibrium,

$$f^e = \frac{(1-\gamma)c}{(1-r)\phi_1}, \tag{13}$$

and

$$v^e = \frac{p\,(1-m)\,k}{(1-p)\,(1-r)\,[\phi_1 - k]} \tag{14}$$

where $c = (C_H - C_L)$.

Before moving to the third party's behaviour, the patient's expected utility (EU) can now be defined as :

$$EU = (y-\alpha) + \left[ph_L^L + (1-p)\left((1-v^e)h_H^H + v^e h_H^L\right)\right] + f^e\left[-p(1-m)k + (1-p)\,v^e\,(1-r)\,(\phi_1 - k)\right] \tag{15}$$

where $h_L^L = h(\theta_L, t_L)$, $h_H^L = h(\theta_H, t_L)$ and $h_H^H = h(\theta_H, t_H)$. Equation (15) states that the patient's expected utility is the sum of three elements. The first term $(y - \alpha)$ represents the income available for consumption goods after paying health insurance premium ; the second $\left[ph_L^L + (1-p)\left((1-v^e)h_H^H + v^e h_H^L\right)\right]$ represents the patient's expected health benefit, and finally $f^e\left[-p(1-m)k + (1-p)\,v^e\,(1-r)\,(\phi_1 - k)\right]$ is the patient's expected net gain from suing a physician given that he observes that the physician chooses t_L.[38]

[38] I showed that if $\theta = \theta_L$, the physician will always behave appropriately, hence the patient's expected health benefit if $\theta = \theta_L$ is ph_L^L. If however $\theta = \theta_H$, I showed that the physician will behave appropriately with probability $(1-v^e)$, thus the patient's expected health benefit if $\theta = \theta_H$ is $(1-p)\left((1-v^e)h_H^H + v^e h_H^L\right)$.

The physician's expected utility (EV) is defined as :

$$\begin{aligned} EV &= \{\rho + \gamma\left[pC_L + (1-p)((1-v^e)C_H + v^e C_L)\right]\} - \\ &\quad \{\left[pC_L + (1-p)\left((1-v^e)\, C_H + v^e C_L\right)\right] + (1-p)v^e f^e \,(1-r)\,\phi_1\} \end{aligned} \tag{16}$$

which can be rewritten as :

$$EV = \rho - (1-\gamma)\left[pC_L + (1-p)\left((1-v^e)\, C_H + v^e C_L\right)\right] - (1-p)v^e f^e\,(1-r)\,\phi_1. \tag{16$'$}$$

Equation (16) states that the physician's expected utility is the difference between the expected reimbursement and the expected costs. The first is the sum of the fixed payment ρ and a share γ of the expected costs of the treatment ; the latter is the sum of the expected costs of the treatment and the expected fine paid by the physician given that he behaved inappropriately and the patient sued.[39]

3.4 The Third Party's Behaviour

As noted above, the health insurer chooses a fixed fee (ρ) and a share (γ) of the treatment cost reported by the physician. Given that the third party is assumed to operate in a competitive market, thus the patient's premium α equals the insurer's expected reimbursement to the physician. Formally,

$$\alpha = \rho + \gamma\left[pC_L + (1-p)((1-v^e)C_H + v^e C_L)\right] \tag{17}$$

If I assume that, at the optimum, the physician's participation constraint is binding,

As a consequence, the patient's expected health benefit is $\left[ph_L^L + (1-p)\left((1-v^e)h_H^H + v^e h_H^L\right)\right]$.

[39] Given that if $\theta = \theta_L$, the physician will always behave appropriately, then the expected cost of the treatment if $\theta = \theta_L$ is pC_L. If however $\theta = \theta_H$, then the expected cost of the treatment is $(1-p)\left((1-v^e)\, C_H + v^e C_L\right)$. Hence the expected cost of the treatment is : $\left[pC_L + (1-p)\left((1-v^e)\, C_H + v^e C_L\right)\right]$.

then α can be rewritten as :

$$\alpha = [pC_L + (1-p)\,C_H] - (1-p)v^e\,[c - f^e(1-r)\phi_1]\,. \qquad (17')$$

3.5 The Court's Behaviour

Given the above results, the court's problem can now be written formally.

The court chooses ϕ_1 to maximize the patient's expected utility subject to (17′). It can easily be shown using the above results that the court's objective function can be rewritten as :

$$\begin{aligned} EU \quad = \quad & y + \left[ph_L^L + (1-p)\,h_H^H\right] - [pC_L + (1-p)\,C_H] + \\ & (1-p)v^e\left[(c - (s-r)(\bar{h} - \widetilde{h})\right] - f^e k\,[p(1-m) + (1-p)\,v^e(1-r)] \end{aligned} \qquad (18)$$

where $\bar{h}$ and $\widetilde{h}$ denote the health benefit associated with a good and a bad outcomes (given that $\theta = \theta_H$), respectively.

The first-order conditions (F.O.C.) of maximization with respect to ϕ_1 gives (for the proof, see the Appendix C) :

$$\frac{p\,(1-m)\,k}{(1-r)}\left[(s-r)(\bar{h} - \widetilde{h}) - \gamma c\right] = 0$$

and consequently[40] :

$$\gamma^* = \frac{(s-r)(\bar{h} - \widetilde{h})}{c} \qquad (19)$$

and

$$\rho^* = (1-\gamma^*)\,[pC_L + (1-p)C_H] \qquad (20)$$

Substituting (13) and (14) into (17′), γ follows as :

$$\gamma = \frac{(1-r)\,(\phi_1 - k)}{p\,(1-m)\,k}\frac{X}{c} \qquad (21)$$

[40] Since p, m and r are non-zero probabilities, it follows that $\frac{p(1-m)k}{(1-r)}$ cannot be equal to zero.

where $X = [(pC_L + (1-p)\, C_H) - \alpha]$

Using (19) and (20), the optimal fine is given by :

$$\phi_1^* = k + \frac{p\,(1-m)\,k\left[(s-r)(\bar{h}-\widetilde{h})\right]}{(1-r)\,X} \tag{22}$$

Given that conditional on $\theta = \theta_H$ the probability that t_H will induce the good state is higher than the probability that t_L will induce it (i.e., $s > r$), it follows that $(s-r)(\bar{h}-\widetilde{h}) > 0$. We know from (17′) that $\alpha = pC_L + (1-p)\, C_H$ if and only if $v^e = 0$, i.e., if and only if the physician will never cheat. Given that we showed that the physician uses a mixed strategy (i.e., $v^e \in (0,1)$), thus α cannot be equal to $pC_L + (1-p)\, C_H$. More specifically, $\alpha < pC_L + (1-p)\, C_H$ (or equivalently $X > 0$), that is, the health insurance premium paid by the patient is less than the insurance premium which prevails in a competitive market with no cheating possibility. This result comes from the fact that the third party knows that, for a patient who has a high-severity illness, the physician will not provide the appropriate treatment with certainty. And, since the third party operates in a competitive market, it must adjust the level of the insurance premium paid by the patient. Since $(s-r)(\bar{h}-\widetilde{h}) > 0$ and $X > 0$, it follows that $\phi_1^* > k$. From (22), $\frac{\partial \phi_1^*}{\partial(\bar{h}-\widetilde{h})} > 0$, i.e., the higher is the health damage, the higher is the compensation.

Using (13), (14) and (22), the physician and the patient's optimal strategies are respectively :

$$v^* = \frac{X}{(1-p)\left[(s-r)(\bar{h}-\widetilde{h})\right]} \tag{23}$$

$$f^* = \frac{(1-r)\left[c-(s-r)(\bar{h}-\widetilde{h})\right] X}{(1-r)^2 k + p(1-m)k\left[(s-r)(\bar{h}-\widetilde{h})\right]} \tag{24}$$

We know by assumption that $\gamma < 1$. Using this assumption, from (19) $(s-r)(\bar{h}-\widetilde{h}) < c$ and consequently it can be shown that $\frac{\partial f^*}{\partial k} < 0$. That is, *ceteris paribus* the higher is the suing costs, the lower is the probability that the patient will sue.

To sum up, the set of equations 19-24 characterizes the equilibrium behaviour of the agents in the model. The court who acts as a regulator of health care sets the compensation higher than the cost of litigation. The third party, given the level of compensation set by the court, and anticipating the physician's reaction about the level of the optimal fine, chooses to remunerate the physician by combining a fixed and positive payment (ρ^*) with a partial reimbursement (γ^*) for the cost of the treatment. The level of the fixed payment chosen by the third party is a fraction of the expected cost of the treatment if we were in a setting where the physician always behaves appropriately. That is, the physician always provides the inexpensive treatment given that the patient has a low-severity illness and the expensive one to a patient who has a high-severity illness. The contract is such that at the equilibrium, given that the third party operates in a competitive market, its expected profits are zero. The physician, given the reimbursement contract, randomizes between treating appropriately or not. More specifically, the physician will never treat appropriately with probability one, as he will also never cheat with certainty. At the equilibrium, the physician's expected utility is equal to his reservation utility which is normalized to zero. Like the physician, the patient randomizes between suing or not. But unlike the physician, at the equilibrium, his expected utility is positive.

The model also predicts that the fees incurred by the patient to sue a physician are negatively related to the suing probability (i.e.,$\frac{\partial f^*}{\partial k} < 0$). More specifically, a higher fee deters the likelihood of filing a lawsuit. Finally, the model predicts that the patient's compensation is proportionate to the harm (i.e., $\frac{\partial \phi_1^*}{\partial (\bar{h} - \tilde{h})} > 0$).

In the next section, I will compare a system with a mixed payment mechanism (which is considered in this model) with that of a pure capitation ($\gamma = 0$) and a pure FFS ($\gamma = 1$).

3.6 The Patient's Expected Utility Under Different Payment Mechanisms

Given the above results for the mixed payment mechanism (i.e., $\gamma < 1$), I can now compare the patient's expected utility in different payment systems. I consider a pure capitation and a pure FFS system, respectively.

3.6.1 The Capitation System

If the partial reimbursement $\gamma = 0$, then from (13), the suing probability $f^* = \frac{c}{(1-r)\phi_1^*} = f^C$. In a mixed payment setting, $f^* = \frac{(1-\gamma)c}{(1-r)\phi_1^*} = f^M$. By comparing f^* in the two systems, it follows that for a given level of ϕ_1^*, $f^C > f^M$. That is, for the same level of compensation, the probability that the patient will sue the physician is higher under the capitation system than the mixed one. Combining (18) with the fact $f^C > f^M$, it can be shown that $EU^C < EU^M$, where EU^C and EU^M denote the patient's expected utility in a capitation and a mixed payment mechanism, respectively. The last inequality comes from the fact that, from (23), the probability that the physician will provide the inexpensive treatment to a patient who has a high-severity illness is the same in the capitation and the mixed payment setting.

3.6.2 The FFS System

I showed that if $\gamma < 1$ (i.e., if the third party reimburses only partially the physician for the cost of the treatment), then $(s-r)(\bar{h}-\widetilde{h}) < c$. If $\gamma = 1$, then from (19), $(s-r)(\bar{h}-\widetilde{h}) = c$. Using this equality, it follows (from (24)) that $f^* = 0 = f^{FFS}$. That is, in a FFS payment mechanism, the patient will never sue. It is important to note that this result comes from the fact that, in the present model bad outcomes are not the result of ignorance but rather strategic behaviour. And, since all health care costs reported by the physician are covered by the health insurance provider, there is no incentives to choose the inappropriate treatment.

The patient's expected utility EU^{FFS} in this setting is given by :

$$EU^{FFS} = y + \left[ph_L^L + (1-p)\, h_H^H\right] - \left[pC_L + (1-p)\, C_H\right] \tag{25}$$

Equation (25) is obtained by substituting the equality $(s-r)(\bar{h}-\widetilde{h}) = c$ and $f^* = 0$ into (18). By examining (18), it is easy to see that EU^{FFS} is equal to the patient's expected utility when $f^* = 0$ and $v^* = 0$. That is, in a FFS payment setting, the patient's expected utility is the same as in the case where the physician always behave appropriately, i.e., the physician never under-treat his patient. Given that EU^{FFS} is $MaxEU$ if $f^* = 0$ and $v^* = 0$, it follows that $EU^{FFS} > EU^M$, i.e, the patient's expected utility is higher in the FFS mechanism than in the mixed system.

To sum up, $EU^C < EU^M$ and $EU^{FFS} > EU^M$ thus $EU^C < EU^M < EU^{FFS}$. Hence, the patient's expected utility is higher in the mixed payment mechanism than in the capitation system, but is lower than the expected utility in the FFS mechanism. These results come from the fact that the physician will : (i) more often under-treat his patient in the capitaion system than in the mixed system and (ii) never under-treat in the FFS payment.

In the next section, I relax the assumption that the court can perfectly observe (ex post) the patient's true illness severity. More specifically, I assume that the court may err with some positive probability and I analyze briefly the implications of this assumption on the patient's suing decision and the physician's treatment choice decision.

4 The Equilibrium with Court Errors

I assume that, given that the patient's true illness severity is θ_L, the court observes θ_H with probability π_1 and θ_L with probability $(1-\pi_1)$. If the true severity is θ_H, I assume that the court observes θ_L with probability π_2 and θ_H with probability $(1-\pi_2)$.

If the patient observes that t_H is provided, again he will never sue. This is because there is no fine for over-treatment and thus even if the physician treats θ_L with t_H, it is better for the patient to not sue.[41] Also, if the true illness severity were θ_H, the physician can not be fined. Now, I will reexamine the physician's treatment choice decision given the type of severity observed. If the physician observes θ_L, he will choose t_H if and only if :

$$f^e > \frac{(1-\gamma)c}{(1-m)\pi_1\phi_1} = \bar{f}^1 \tag{26}$$

If the physician observes θ_H, he will choose t_H if and only if :

$$f^e > \frac{(1-\gamma)c}{(1-r)(1-\pi_2)\phi_1} = \bar{f}^2 \tag{27}$$

where $\bar{f}^1$ ($\bar{f}^2$) denotes the suing probability which makes the physician indifferent between treating appropriately or not, given that the patient has the low-severity illness (the high-severity illness).

Equations (26) and (27) follow from the fact that if the physician chooses t_H, his expected utility is (as before) $\rho - (1-\gamma)C_H$. But if he chooses t_L, his expected utility becomes $\rho - (1-\gamma)C_L - (1-m)f\pi_1\phi_1$ if $\theta = \theta_L$ and $\rho - (1-\gamma)C_L - (1-r)f(1-\pi_2)\phi_1$ if $\theta = \theta_H$.

Remember that it was shown above that at the equilibrium, $f^e = \bar{f}$ (see equation (13)). Similarly, at the equilibrium, $f^e = \bar{f}^1$ if $\theta = \theta_L$ and $f^e = \bar{f}^2$ if $\theta = \theta_H$. Given that $m > r$, thus $\bar{f}^1 > \bar{f}$. Also, since π_2 is a non-negative probability, it follows that $\bar{f}^2 > \bar{f}$. Hence, if the court can mistakenly fine a physician who treats appropriately his patient, then at the equilibrium, the probability that the patient will sue a physician is higher than in a case where the court's decision is perfect. The intuition behind this result is simple : the patient, knowing that the court can err, will be more likely to sue the physician even if he believes that the appropriate treatment was provided. Since the physician's behaviour when θ_L is observed is based on $\bar{f}^1$, thus in comparison with the case without court errors, here the

[41] If he sues, he will lose k.

physician will treat a patient who has a low-severity illness with the expensive treatment, i.e., over-treatment may occur at the equilibrium. Hence, the fact that the physician can be fined mistakenly induces the latter to practice defensive medicine.

In sum, when the court's decisions are imperfect, the suing rate will be high and some physicians who treat appropriately their patients may be sued. Physicians knowing this, will over-treat their patients with some probability, an equilibrium which can not occur in the perfect information case.

In the next section, I relax the assumption that the patient is risk neutral and given the parameters of the model, I analyze the risk-averse patient's suing behaviour and compare it with the suing behaviour of a risk-neutral law firm.

5 The Risk-Averse Patient vs the Risk-Neutral Law Firm

In this section I examine two cases where a risk-averse patient : (i) must assume all suing responsibilities and costs or (ii) the suing responsibilities are transferred to a law firm which covers all lawsuit costs in exchange for a given proportion of the compensation if the lawsuit is succesful and zero if not. For simplicity, let us denote by q the probability of wining the medical malpractice claim with ϕ as the corresponding award. Also, without loss of generality I assume that the patient's utility function depends on the individual's monetary wealth.[42] More specifically,

$$U^P = \begin{cases} U(y-\alpha) \text{ if the patient does not sue} \\ U(y-\alpha+\phi-k) \text{ if the patient sues and win} \\ U(y-\alpha-k) \text{ if the patient sues but loses} \end{cases}$$

where $y - \alpha$ is the income available for consumption goods after paying health insurance premium α, and k is the lawsuit cost (as defined previously). In the next sections, I analyze a suing decision for a risk-averse patient and a risk-neutral law firm, respectively.

[42] Ignoring the health state component in the utility function does not affect in any way the results.

5.1 The Risk-Averse Patient

We know that a risk-neutral patient will sue if and only if the expected utility of suing is greater than or equal to the utility of not suing. Formally, the patient will sue if and only if :

$$q\left[U(y-\alpha+\phi-k)\right]+(1-q)\left[U(y-\alpha-k)\right]\geq U(y-\alpha) \tag{28}$$

Since the patient is risk neutral, equation (28) can be rewritten as :

$$q(y-\alpha+\phi-k)+(1-q)(y-\alpha-k)\geq(y-\alpha)$$

which reduces to :

$$\phi\geq\frac{k}{q}=\bar{\phi}^{P} \tag{29}$$

where $\bar{\phi}^{P}$ denotes a risk-neutral patient's threshold of indifference between suing or not. In other words, equation (29) states that a risk-neutral patient's decision is as follows : sue if $\phi>\bar{\phi}^{P}$; do not sue if $\phi<\bar{\phi}^{P}$. Finally the patient is indifferent between suing and not if $\phi=\bar{\phi}^{P}$.

If however the patient is risk averse, his suing strategy is different and depends on his level of aversion. In fact, for a risk-averse patient, the utility function is concave, and given as :

$$q\left[U(y-\alpha+\phi-k)\right]+(1-q)\left[U(y-\alpha-k)\right]<U\left[q(y-\alpha+\phi-k)+(1-q)(y-\alpha-k)\right] \tag{30}$$

that is, the expected utility is lower than the utility of the expected gain. If $\phi=\frac{k}{q}$ then $U\left[q(y-\alpha+\phi-k)+(1-q)(y-\alpha-k)\right]=U(y-\alpha)$ i.e., the utility of the expected gain is equal to the utility of the initial wealth (or equivalently, the expected gain is equal to the initial wealth). As a consequence, (30) can be rewritten as :

$$q\,[U(y-\alpha+\phi-k)]+(1-q)\,[U(y-\alpha-k)]<U(y-\alpha) \tag{30$'$}$$

We know that a patient will sue if equation (28) is satisfied. Hence, comparing (28) and (30$'$), it is easy to see that at the risk-neutral patient's threshold of indifference between suing or not (i.e., at $\bar{\phi}^P = \frac{k}{q}$), the risk-averse patient will not sue. In other words, the risk-averse patient's threshold of indifference between suing or not is higher in terms of the award than that of the risk-neutral patient. The fact that the patient knows that he must pay the lawsuit costs whether he wins or not, serves as a disincentive, especially when the expected award is relatively small. As a consequence, at the equilibrium, the claiming rate will be low and some risk-averse patients who are treated inappropriately will not sue.

In the next section, I end by analyzing the case in which the patient makes an arrangement with a law firm and the latter undertakes all responsibilities with respect to the suing decision.

5.2 The Risk-Neutral Law Firm

I now assume that the risk-averse patient contracts with a law firm who decides either to sue or not. The law firm is responsible for the lawsuit costs and receives a proportion λ of the patient's compensation if the case against the physician is sucessful with the patient receiving $(1-\lambda)$. If however, the law firm loses, the patient receives nothing. This type of agreement between a law firm and a patient (more broadly a client) where the fee is only charged when the lawsuit is successful is termed in legal literature as *contingency fee*. Most contingency fees can range from one-third to 50% of the award.[43] This high percentage can be explained by the fact that, given that the law firm has many cases to defend, the

[43] Source : 'Medical Malpractice Law in the United States', Kaiser Family Foundation. May 2005.

expected gain when it wins must compensate the expected costs in cases lost.

Formally, for a given level of compensation ϕ, the law firm will sue if and only if : $\lambda q\phi - k \geq 0$, that is, $\phi \geq \frac{k}{\lambda q} = \bar{\phi}^L$, where $\bar{\phi}^L$denotes a risk-neutral law firm's threshold of indifference between suing or not. For the patient, if the law firm sues, his expected utility is $q\left[U(y-\alpha+(1-\lambda)\phi)\right]+(1-q)\left[U(y-\alpha)\right]$ which is always greater than $U(y-\alpha)$ if $\lambda \leq 1$. Since the optimal agreement between the patient and the law firm must be such that $\lambda^* \leq 1$, thus a risk-averse patient will always prefer to transfer the suing responsibilities to a law firm. But, as I showed above, the law firm's suing decision depends on the value of his threshold of indifference ($\bar{\phi}^L$) and thus it will not always accept to take on a case. Given that $\lambda^* \leq 1$, then it follows that a (risk-neutral) patient's threshold of indifference between suing and not (i.e., $\bar{\phi}^P$) is lower than that of the law firm (i.e., $\bar{\phi}^L$). In other words, the law firm will accept the contingency fee's agreement if and only if ϕ is sufficiently high. More specifically, if $\phi \leq \bar{\phi}^P$, then the risk-averse patient will never sue ; if $\phi \in (\bar{\phi}^P, \bar{\phi}^L)$, then the patient may sue ; finally if $\phi \geq \bar{\phi}^L$, then the law firm will sue.[44]

To sum up, the patient's expected utility is always higher if a third party such as a law firm assumes all the suing responsibilities. However, because the law firm receives only a proportion of the award, it will accept the contingency fee contract if and only if the award is relatively high. Moreover, given that the level of the patient's compensation is proportional to the harm, the law firm will only accept cases involving a serious injury. It follows that, if the amount of compensation is reduced, then the lawyers will refuse some cases.

---

[44] As a numerical example, assume that the probability to win a medical malpractice claim against a physician is 50% and if the latter is won, each party receives 50%. The litigation costs are assumed to be \$10,000. It follows that the patient's threshold of indifference is \$20,000 (i.e., $\frac{10000}{0,5}$) and that of the law firm is \$40,000 (i.e., $\frac{10000}{0,5*0,5}$). As a consequence, if the award is lower than \$20,000, then both the patient and the law firm will not sue. If the award is higher than \$20,000 but lower than \$40,000, then the patient may sue but the law firm will not. Finally, if the award is higher than \$40,000, then the law firm will sue.

6 Conclusion

In this paper, I examine a model in which both physician payment and medical liability mechanisms may serve as a way to provide efficiently health services. To this end, I derive simultaneously in a game-theoretical model, optimal payment contracts for the physician and optimal compensation for the patient given that he is injured by the physician's action. The main contribution of the model proposed here is that both physician payment and medical liability mechanisms are endogenized.

Initially, by assuming that all agents in the model are risk neutral, I find that, at the equilibrium, the physician randomizes between treating appropriately or not and as a consequence the patient also randomizes between suing or not. I also find that the possibility of court errors induces some physicians to practice defensive medicine. That is, over-treatment may occur at the quilibrium. The model also predicts findings that are consistent with the real world observations. First, it predicts that the fees incurred by the patient to sue a physician is negatively related to the suing probability ; and second, the patient's compensation is proportional to the damage.

By relaxing the risk neutrality assumption on patients' side, I find that some patients who are treated inappropriately will not sue and as a consequence, the claiming rate will be low relative to efficient levels. I also find that a risk-averse patient will always prefer to transfer the suing responsibilities to a law firm ; but the latter will only accept cases involving a serious injury. Finally, the model also predicts that policies that tend to impose caps on medical malpractice awards will lower the claiming rate ; this is because many cases will become unattractive for lawyers who accept the contingency fee contract.

7 References

[1] Arlen Jennifer and MacLeod Bentley (2005) : 'Torts, Expertise and Authority : Liability of Physicians and Managed Care Organisation', *Rand Journal of Economy* 36, 494-519.

[2] Becker Gary (1968) : 'Crime and punishment : an economic approach', *Journal of Political Economy* 76, 169-217.

[3] Ake Blomqvist (1991) : 'The doctor as double agent : information asymmetry, health insurance, and medical care', *Journal of Health Economics* 10, 411-432.

[4] Danzon Patricia (1985) : 'Liability and liability insurance for medical malpractice', *Journal of Health Economic*s 4, 309-331.

[5] Danzon Patricia (2000) : 'Liability for medical malpractice'in *Handbook of Health Economic*s.

[6] Ellis P. Randall and McGuire Thomas (1986) : 'Provider behavior under prospective reimbursement', *Journal of Health Economics* 5, 129-151.

[7] Ellis P. Randall and McGuire Thomas (1990) : 'Optimal payment systems for health services', *Journal of Health Economics* 9, 375-396.

[8] Gal-Or Esther (1999) 'Optimal Reimbursement and Malpractice Sharing Rules in Health Care Markets ', *Journal of Regulatory Economics* 16, 237-265.

[9] Léger Pierre Thomas (2000) 'Quality control mechanisms under capitation payment for medical services', *Canadian Journal of Economics* 33, 564-88.

[10] Ma Ching-to Albert and McGuire Thomas (1997) : 'Optimal health insurance and provider payment', *American Economic Review* 87, 685-704.

[11] Shavell Steven (1980) : 'Strict liability versus negligence', *Journal of Legal Studies* 9, 1-25.

[12] Shavell Steven (1991) : 'Specific versus general enforcement of law', *Journal of Political Economy* 99, 1088-1108.

[13] Selden Thomas (1990) : 'Amodel of capitation', *Journal of Health Economics* 9, 397-410.

[14] Zeiler (2008) : 'Contract disclosure rules and damages caps : an equilibrium model of their effects on behavior in health care markets', *working paper.*

8 Appendix

8.1 Appendix C : Proof of ϕ_1^*.

$$EU = A + (1-p)v^e\left[(c-(s-r)(\bar{h}-\widetilde{h})\right] - f^e k\left[p(1-m)+(1-p)\,v^e(1-r)\right]$$

where $A = y + \left[ph_L^L + (1-p)\,h_H^H\right] - \left[pC_L + (1-p)\,C_H\right]$. Substituting for v^e and f^e, EU becomes :

$$\begin{aligned} EU &= A + \frac{p\,(1-m)\,k}{(1-r)\,(\phi_1-k)}\left[(c-(s-r)(\bar{h}-\widetilde{h})\right] - \\ &\quad \frac{(1-\gamma)c}{(1-r)\phi_1}k\left[p(1-m)+\frac{p\,(1-m)\,k}{(\phi_1-k)}\right] \end{aligned}$$

EU can be rewritten as :

$$\begin{aligned} EU &= A + \frac{B}{(\phi_1-k)} - \frac{C}{\phi_1}\left[D+\frac{Dk}{(\phi_1-k)}\right] \\ &= A + \frac{B}{(\phi_1-k)} - \frac{C}{\phi_1}\left[\frac{D\phi_1}{(\phi_1-k)}\right] \\ &= A + \frac{B-CD}{(\phi_1-k)} \end{aligned}$$

where $B = \frac{p(1-m)k}{(1-r)}\left[(c-(s-r)(\bar{h}-\widetilde{h})\right]$, $C = \frac{(1-\gamma)c}{(1-r)}k$ and $D = p(1-m)$.

$$\frac{\partial EU}{\partial \phi_1} = \frac{-(B-CD)}{(\phi_1-k)^2} = 0$$

which is equivalent to :

$$(B-CD) = 0$$

Finally, substituting for B, C and D, we have :

$$\frac{p\,(1-m)\,k}{(1-r)}\left[(s-r)(\bar{h}-\widetilde{h}) - \gamma c\right] = 0$$

Essay 3
Variations in Obstretricians' Use of C-Sections : Evidence from Canada

Abstract

In this paper, I investigate whether and to what extent the variations in c-section rates are due to physician factors. More specifically, using data on deliveries in Canada between 2005 and 2009, I examine the effect of physician factors on the physican's decision to perform a c-section. Controlling for observable patient characteristics, I find that physician factors are an important determinant in the c-section decision and that there is substantial variation across physicians in c-section rates within the same hospital. Thus, the results suggest that some patients are receiving c-sections simply because of their physician's type.

1 Introduction

Many studies have shown that the type and amount of treatment a patient receives depend not only on patients' medical conditions but also on where patients live and that there is substantial variation in the practice of medicine.[45] As a consequence, some patients may receive too little or too much treatment. Practice variations exist and are well documented in many areas of health services.[46] These practice variations may lead to welfare losses if health care professionals deviate from cost-effective treatments. In fact, Phelps and Parente (1990) estimated that the annual welfare loss in 1987 due to hospital rate variations in the top 25 medical interventions in United States exceeds \$7 billion.[47]

Besides physician payment mechanisms and patient characteristics (such as medical conditions, preferences and socioeconomic status), several other factors have been identified as sources of practice variations. These factors include physician characteristics such as age, experience, skills and training (Phelps, (2000) ; Grytten and Sorensen (2003) and Epstein and Nicholson, (2009)) and organizational factors such as hospital processes, norms and standards (Groenewegen and Westert, (2004) and Bynum, Song and Fisher, (2010)). The objective of this paper is to examine whether and to what extent physician factors influence practice variations in cesarean section (c-section) surgeries among obstetricians practicing in the same hospital.

C-section rates are extremely high in most developed countries and are beyond the maximum level of 15% recommended by the World Health Organization (WHO) in 1985, which

[45] See Wennberg and Gittelsohn, (1973) ; Wennberg, Fisher and Skinner, (2002) and Fisher, Bynum and Skinner, (2009).

[46] See for example, McMahon et al., (1999) and Bynum, Song and Fisher, (2010) for practice variations in cancer diagnosis and treatments ; Grant, (2005) and Epstein and Nicholson, (2009) for c-section surgeries ; Hanan, Wu and Chassin, (2006) and Tu et al., (2012) for cardiovascular care.

[47] These interventions include hospitalizations for coronary artery bypass graft, psychosis, cardiac catheterization and adult pneumonia.

represents the clinically justified rate.[48] In Canada, for example, the Canadian Institute for Health Information (CIHI) reports that the c-section is one of the most common surgical procedures performed on Canadian women and that the c-section rate was about 28% in 2008, which is nearly twice the WHO's justified rate. It also reports that, compared with vaginal births, c-section deliveries can cost hospitals twice as much in obstetric care for both mothers and babies. Also, in Canada, most of obstetrician/gynecologists are paid on a fee-for-service basis and this may lead them to choose the c-section delivery even if the vaginal delivery is more appropriate given the patient's observable characteristics.

Clinical and socio-demographic factors such as the underlying medical conditions of the mother, complications related to pregnancies and maternal age are often cited as determinants of c-sections. However, it has been shown that these factors do not fully explain the excessively high c-section rate and that nonclinical factors may also lead physicians to choose the c-section delivery.

Among these nonclinical factors studied (or identified) are : physicians' fears of malpractice lawsuits (Localio et al., (1993) ; Baldwin et al., (1995) ; Sloan et al., (1997) and Dubay et al.,(1999)), physician financial incentives (Gruber and Owings, (1996) ; Gruber et al., (1999) and Grant, (2009)) and physician-specific characteristics (Tussing and Wojtowycz, (1992) ; Burns et al., (1995) ; Horwitz et al., (2000) ; Grant, (2005) and Epstein and Nicholson, (2009)).

Although it is well documented that there is substantial variation across physicians in their propensity to perform a cesarean section, controlling for patients' observed characteristics, previous estimates may be biased because of self-selection or endogeneity problems. In fact, some mothers may select their obstetrician according to their preferences for a c-section and also some physicians may perform high numbers of deliveries for mothers with

[48] The clinically justified rate refers to the rate which will prevail if every physician follows clinical practice guidelines.

risk factors. That is, previous studies do not explicitly consider the problem of patient selection of physicians on the basis of unobserved characteristics of mothers, or the fact that some physicians are specialized in the treatment of mothers with specific health conditions (Grant, (2005) and Epstein and Nicholson, (2009)), or relies on a sample that may contain nonrandomly paired patients and physicians (Epstein, Ketcham and Nicholson, (2010)). In this paper, I also examine the extent to which variations in c-section rates are due to physician factors but using institutional level data where patients are more likely to be randomly assigned to physicians. To this end, I use the Discharge Abstract Database (DAD) of the CIHI. In the data set, the admission category for obstetric patients can be either elective or urgent/emergent. An admission is coded as elective if the hospital or the facility is expecting the patient (i.e., it includes patients booked for elective c-section deliveries and those who are at term and are pre-booked) while an urgent or emergent admission mostly refers to cases where the patient is not expected in the facility. By focusing only on urgent admissions, one can deal with or at least limit the potential selection bias (i.e., the problem of patient selection of physicians on the basis of unobserved characteristics of the mothers). This is because, given that the hospital or the facility is not expecting these patients, they are likely to be assigned to other physicians rather than to their own obstetricians.

As noted above, this paper is related to Grant (2005) and Epstein and Nicholson (2009) who analyze variations across physicians in c-section rates, controlling for patients' observed characteristics. These papers find substantial variations across obstetricians. More specifically, the first paper finds a standard deviation of 6.6 percentage points, using Florida's hospital discharge data for the year 1992. Similarly, the second paper finds a standard deviation of about 5 percentage points with panel data (1992-2006) on deliveries in Florida and New York states. However, these papers remain silent on the endogeneity problem. These papers also do not explicitly consider the level of these variations which are due to physician practice styles. That is, they cannot (or do not) disentangle the separate effect of physician treatment styles and hospital characteristics on the variations across physicians in c-section rates.

Finally, the paper is also related to Epstein, Ketcham and Nicholson (2010) who examine the variation in physicians' skill at performing c-sections and whether patients are appropriately matched to physicians in obstetric markets. The authors explicitly consider the problem of patient selection of physicians by using the random pairing of physicians and patients on weekends to measure obstetricians' treatment styles. More specifically, by assuming that patients who go into labor on a weekend are randomly assigned to an on-call obstetrician, they are able to measure more precisely the effect of physicians' treatment styles. Although their approach to deal with self-selection issues is novel, the randomness of their sample is somewhat questionnable. In fact, in practice, many hospitals or facilities do not randomize obstetricians on weekends ; and also as admitted by the authors, the weekend sample may not be completely random because they do not observe whether the delivery was induced or not.

As noted above, the present paper focuses on urgent admissions to control for self-selection issues. Although I rely on this specific framework. I also perform a test of the randomness of patients to physicians to determine whether or not the proportion of patients with observable risk factors is statistically similar across each physician's group of patients. Also, in order to disentangle the effect of physician factors from hospital characteristics (such as hospital processes, norms and standards) on the variations across physicians in c-section rates, I focus on within-hospital variations. Doing so allows us to perfectly isolate the level of variations in risk-adjusted cesarean section rates which is due to physician factors. That is, variations net of all hospital level effects including unobserved hospital-level effects such as the population it attracts (i.e., conditional on the patient population).

Results show that physician factors are an important element in the c-section decision and omitting them can lead to biased estimates by imputing some of the variation in c-section rates to factors other than physician characteristics. The results also suggest that there is substantial variation across physicians in c-section rates, and in many hospitals these variations are relatively higher than those found by Grant (2005) and Epstein and Nicholson (2009). More specifically, I find that, in most hospitals, the delivery method (c-section vs. vaginal birth) depends a lot on to whom you have been assigned. As a consequence, if the average risk-adjusted c-section in each hospital is considered as the appropriate treatment given the hospital's characteristics, then the results suggest that there will be welfare losses due to these variations.

The remainder of the paper is organized as follows. In section 2, I present the data and descriptive statistics. Sections 3 and 4 present the randomness test and the empirical model, respectively. In section 5, the results of the estimations are provided. Finally, section 6 concludes.

2 The Data and Summary Statistics

In this paper, I use the Discharge Abstract Database (DAD) from CIHI, which contains information on all deliveries across Canada, excluding the province of Quebec. For every delivery, the database includes the mother and newborn characteristics (such as mother's age, gestational age, birth weight,...), detailed information on up to 25 diagnosis codes and all procedure codes, separate identifiers for the hospital and the physician, and the province or territory of practice. Because each physician has a unique identifier, the physician's practice can be followed over time. But, given that each hospital assigns a unique identifier to the physician, we cannot follow physicians when they change hospitals.[49] It is important

[49] As a consequence, the analysis can only be done at hospital level.

to note that in the data set the only information about the physician is his identifier. That is, there is no information on each physician characteristics such as age, experience, gender and medical training. Furthermore, because the mother's chart number is recorded on all newborn abstracts and the newborn's chart number is also recorded on the mother's abstract, we are able to merge the maternal and newborn records.

The procedure and diagnostics codes allow the researcher to observe whether the delivery is by c-section or vaginally and us to define the variables which influence the physician's decision to perform a c-section such as the health conditions of the mother and the newborn. These variables include : fetal distress (low oxygen levels in the fetus), dystocia (failure to progress in labor), gestational age of the baby, maternal and gestational hypertension and abnormal presentation of the foetus (any position in which a part of the baby other than the crown of its head emerges first).

The data set contains information on over 1.3 million deliveries performed during the 2005-2009 period. Table 3.I presents the distribution of deliveries and c-section rates over this time period. The number of deliveries increased between 2005 and 2009. The sample consists of 1,317,297 deliveries performed in Canada by 5,889 health care professionals.[50] Of these deliveries, 27.38% were by c-section.

The distribution of all deliveries and c-section rates by province is reported in Table 3.II below. Both the number of deliveries and the c-section rate vary across Canada. Of all deliveries, 650,631 (49%) were performed in Ontario, which is the province with the highest number of deliveries compared to Prince Edward Island which is the province with the lowest one with 6,553 (0.5%) deliveries. Besides Manitoba and Saskatchewan, all provinces have a c-section rate higher than 27%. Newfoundland with more than 30% is the province with the highest c-section rate and Manitoba has a c-section rate of nearly 20% which

[50] Territories (i.e., Northwest Territories, Nunavut and Yukon) are excluded.

represents the lowest provincial rate in the country.

Table 3. I – Number of Deliveries and C-Section Rates by Year, 2005-2009

Year	Number of deliveries	C-section rate
2005	249,519	0.2701
2006	250,085	0.2702
2007	266,119	0.2744
2008	274,628	0.2783
2009	276,946	0.2752
2005-2009	1,317,297	0.2738
Number of physicians	5,889	

Table 3. II – Number of Deliveries and C-Section Rates by Province, 2005-2009

Province	Number of deliveries	C-section rate.
Newfoundland (NFLD)	21,002	0.3050
Prince Edward Island (PEI)	6,553	0.3025
British Columbia (BC)	197,366	0.3000
New Brunswick (NB)	34,371	0.2840
Ontario (ON)	650,631	0.2797
Alberta (ALB)	227,235	0.2709
Nova Scotia (NS)	42,774	0.2692
Saskatchewan (SASK)	64,089	0.2147
Manitoba (MAN)	73,276	0.1975

In order to deal with the potential selection bias, I focus only on urgent admissions. According to the DAD abstracting manual, the deliveries coded (by triage nurses) as urgent include : (i) patients at term, with no pre-natal care and who are not known at the facility, (ii) patients where the mother's or undelivered baby's life is threatened, (iii) patients admitted postpartum and who had not been pre-booked, (iv) pre-term patients (less than 37 weeks of gestation) and (v) pre-term patients, pre-booked and who arrive with an urgent medical problem and require immediate delivery.[51] For the analysis, all pre-term patients are excluded. I do this because some of them (i.e., patients in category (v)) are pre-booked

[51] These categories (i.e., emergency admissions) represent 12.1% (158,267) of all deliveries. Of these emergency deliveries, 32.8% were by c-section, which represents an increase of about 25% compared to the rate in the full sample.

and may be matched with their obstetrician and there is no information allowing us to exclude only these patients. It is important to note that, although patients in the first category are randomly assigned to obstetricians, some of the patients in categories (ii) and (iii) may not. Unfortunately, in the data set, it is difficult to determine whether or not they are assigned to their own obstetrician. Nonetheless, in the empirical model (detailed in the next section), these situations are taken into account by performing a test of the randomness of patients to physicians. Also, in order to rule out the fact that the variations in c-section rates may be driven in part by women's preferences rather than physician factors, the analysis is restricted to women who went into labor.[52] Using a classification developed by Henry et al., (1995) and Gregory et al., (2002), women who had a vaginal delivery or discharge codes representing fetal distress, labor abnormalities and cord prolapse (the umbilical cord precedes the fetus' exit from the uterus) are categorized as having labored and all other deliveries are classified as scheduled c-sections.

The final sample consists of 83,125 deliveries in 388 hospitals. Table 3.III provides summary statistics for the full sample and for both c-section and vaginal deliveries.

Of all deliveries, about 12% were by c-section.[53] Each physician in our sample has performed on average approximately 24 urgent deliveries during the 2005-2009 period. The majority (more than 80%) of mothers are younger than 35 years old and more than 56% of them are between twenty-five and thirty-five years old. About 54% of all mothers had previous live births.

Among women who had previous pregnancies, the rate of pre-term and term deliveries are 4.3% and 53.5%, respectively, and the spontaneous abortion's rate is about 23%. Pregnancy-related variables (such as multiple gestation, abnormal presentation and gesta-

[52] This restriction excludes 12,938 deliveries (out of 96,063) considered as scheduled c-sections. However, it is worth noting that, these scheduled c-sections may represent many of the more complicated cases and excluding them from the sample may limit a generalization of the findings.

[53] Including scheduled c-sections, the rate is 24.16%, which represents a two-fold increase.

Table 3. III – Sample Means, 2005-2009

Variable	All deliveries	with c-section	without c-section
Delivery method			
C-section	0.1235	1	0
Age groups			
Age<20	0.072	0.059	0.074
Age20-25	0.201	0.172	0.205
Age25-30	0.289	0.273	0.291
Age30-35	0.275	0.295	0.272
Age35-40	0.134	0.161	0.130
Age40-45	0.027	0.037	0.025
Age>45	0.001	0.002	0.001
Previous pregnancies			
Previous live births	0.538	0.246	0.580
Previous term deliveries	0.535	0.241	0.574
Previous pre-term deliveries	0.043	0.243	0.045
Previous spontaneous abortions	0.225	0.208	0.227
Previous c-section	0.035	0.046	0.034
Pregnancy-related variables			
Multiple Gestation	0.002	0.016	0.001
Dystocia	0.118	0.507	0.064
Abnormal presentation	0.025	0.057	0.020
Fetal distress	0.242	0.293	0.663
Premature rupture of membrane	0.114	0.141	0.110
Diabetes	0.037	0.061	0.034
Placenta praevia	0.009	0.031	0.006
Maternal Hypertension	0.007	0.015	0.006
Gestational Hypertension	0.064	0.117	0.056
Macrosomia	0.018	0.034	0.016
Post-term delivery	0.003	0.003	0.003
Deliveries per physician	23.98		
Number of deliveries	83,125	10, 269	72,856

tional hypertension) are used to control for health conditions that may affect the woman's need for a c-section and thus may affect the physican's decision to perform a c-section. For all deliveries, only 0.2% are coded as consisting of two or more babies and 3.5% of mothers had a previous c-section. Table 3.III also shows that about 12% were diagnosed as dystocia, 2.5% as abnormal presentation and 24% as fetal distress. In almost 12% of cases, there was premature rupture of membranes (the membranes break before labor). About 4% of mothers are diagnosed with diabetes and 1% with placenta praevia (abnormal position of the placenta). The maternal hypertension is coded for less than 1% of all deliveries while the gestational one accounts for 7%. Table 3.III also presents some statistics related to the newborn. More specifically, the rate of macrosomia (birth weight greater than 4000 g) is about 2%, and only 0.3% of all deliveries are post-term births (greater than 42 weeks of gestation). Finally, table 3.III also shows that most of the risk factors are more present in births that are delivered by c-section than those delivered traditionally.

3 The Randomness Test

As noted above, some of the patients in the final sample may not have been randomly assigned to their obstetricians and ignoring this problem may lead to biased estimates of the effect of physician factors on the c-section decision. In this section, in order to take into account the problem of patient selection or matching between physicians and patients, I compare, in each hospital, each physician's patient population based on observable patient characteristics. That is, I test (conditional on patient observable characteristics) if some physicians are performing more c-section deliveries simply because they treated high numbers of mothers with preferences for a c-section or with risk factors. To this end, for each patient's characteristic, I use a F-test to see whether or not the patients are significantly different across physicians. In addition to the variables listed in table 3.III above, income quintiles are also included as a measure of socio-economic status because they are likely to affect the c-section decision.

More specifically, for each hospital, I separately estimate the following equation :

$$CX_i = \alpha_j J + \epsilon_{ij} \tag{31}$$

where CX_i is equal to one if the patient's observable characteristic i is present in all deliveries performed in a given hospital, and zero otherwise ; J represents a vector of physician fixed effects and ϵ_{ij} is a standard error term. In equation (31) both the dependent and independent variables are dummy variables. As such, coefficient α_j is easily interpretable as the proportion of patients with characteristic i in the group of patients treated by the physician j. For example, if X_i represents an abnormal presentation, then α_j is the percentage of women with abnormal presentations in physician j's patient pool in a specific hospital. For each patient's characteristic, I test whether or not the coefficients α_j significantly differ across physicians. That is, if some physicians are treating higher numbers of patients with risk factors associated with c-section deliveries.

It is important to note that although equation (31) can be estimated for each hospital (with two or more physicians), the analysis is limited to a select number of hospitals. In fact, in order to rule out the possibility that variations in c-section rates are due to a small number of patients per physician rather than physician factors, I restrict the analysis to physicians who performed at least 50 deliveries in the final sample over the five-year period. This restriction combined with the fact that variations in c-section rates can only be computed for hospitals with at least two physicians excludes the following cases : (i) hospitals with less than 100 deliveries over the entire period (i.e., 259 out of 388 hospitals) and (ii) hospitals with more than 100 deliveries but where only one physician performed more than 50 deliveries (i.e., 84 out of 388 hospitals). As a consequence, the analysis is limited to 45 hospitals (17 in Ontario, 16 in Alberta, 5 in Manitoba, 3 in Saskatchewan, 2 in Nova-Scotia, 1 in Newfoundland and 1 in British Columbia).[54]

The results show that for some observable patient characteristics, the coefficients α_j are not jointly equal to zero. That is, the results suggest that there are some systematic differences across physicians with respect to some patient characteristics. In order to limit the effect of this non-randomness on the likelihood of performing a c-section, these characteristics will be included in the regression as control variables. However, if some of these control variables are endogeneous, then including them may introduce another source of bias. It is also important to note that, among all hospitals in our sample, the results do not support any perfect randomization (based on observable characteristics) assumption and, in some extreme cases (four hospitals), all α_j are not jointly equal to zero for all patient characteristics. However, also noted above, including these variables in the empirical specification may reduce the potential bias due to the non-randomization.

4 The Empirical Specification

In this section, I present an econometric model to estimate the probability that, given the sociodemographic characteristics and clinical indications and physician dummies, a woman will receive a c-section. I consider the following specification :

$$C_{ij} = \beta X_i + \delta_j J + \mu_{ij} \tag{32}$$

where C_{ij} is equal to one if mother i received a c-section by physician j, and zero otherwise ; X_i denotes a vector of observable patient characteristics which are non-random variables (following the test performed in the previous section) ; J represents a vector of physician fixed effects (as defined previously), μ_{ij} is a standard error term. The coefficient δ_j

[54] The number of deliveries, physicians and c-section rates by hospital are reported in table 3.VI in appendix D.

can be interpreted as a physician j' s treatment or delivery style given medical diagnostics and patient characteristics. In other words, δ_j represents a physician j' s risk-adjusted cesarean rate.

Equation (32) is estimated by ordinary least squares rather than a probit or logit model because it is well known in the econometric literature on nonlinear panel data models with unobserved heterogeneity that unconditional logits or probits lead to inconsistent estimators.[55] However, it is worth noting that the Linear probability model is associated with at least two problems. First, it can yield predicted values (of the dependent variable) outside of the 0-1 interval and it is difficult to interpret these values as probabilities ; and second, the estimators are inefficient. The first problem is irrelevant or less important in this paper because the main objective of the paper is to estimate consistently the fixed-effect coefficients. In our estimation, we deal with heteroscedasticity by clustering on physician indicator variables.

In order to test the importance of physician factors in the c-section decision, for physicians in a given hospital, equation (32) is estimated with and without physician fixed effects. In fact, a comparison between the R squared of the two models allows us to determine the percentage of variance in c-section rates that is explained by the variation in physician factors.

In order to examine whether physicians vary in their likelihood to perform a c-section, for each hospital, I estimate equation (32) and I compute the standard deviation of the coefficients δ_j. Formally, I compute $\sqrt{E\left[\hat{\delta}_j - E(\hat{\delta})\right]^2}$, where $E(\hat{\delta})$ represents the mean risk-adjusted cesarean rate in a given hospital. If physicians who practice in the same hospital treat differently patients with similar observable characteristics, then the standard deviations of the risk-adjusted cesarean rates will be relatively high in each hospital.

[55] See for example Neyman and Scott (1948) for more details.

5 The Estimation Results

Although the main objective of this paper is to estimate the coefficients of physician fixed effects and thus to examine the variation across physicians in c-section rates, I begin by presenting the results of the estimated coefficients of patient characteristics. Table 3.IV below reports the results from the estimation of the linear probability model of one hospital.[56]

I select the results of one of the hospitals where patients are likely to be randomly assigned to physicians with respect to many observable patient characteristics. In fact, the results of these hospitals are more likely to display the importance of physician factors in the c-section decision.

First, it is important to note that patient characteristics listed in table 3.III and which are excluded from table 3.IV are those which respect the random matching assumption. As a consequence, one may see (by comparing the two tables) that physicians treat similar patients with respect to many observable patient factors. Among these factors are most of the variables related to previous pregnancies (previous live births, previous c-sections and previous pre-term and term deliveries) and pregnancy-related variables (dystocia, fetal distress, diabetes, gestational hypertension and macrosomia). The results suggest that some observable patient characteristics are important determinants of the c-section decision. For instance, for a woman between the age of 30 and 34, the probability of a cesarean delivery increases by about 5 percentage points compared to a woman who is under 20 years old (which is the reference group). The results also show that the presence of an abnormal presentation significantly increases the probability of having a c-section by approximately 11 percentage points while a post-term delivery decreases the likelihood of a c-section delivery by about 30 percentage points. Finally the results suggest that the patient's socio-economic

[56] All results are available upon request.

Table 3. IV – Estimation Results : Linear Probability Model

Variable	Coefficient	(Std. Err.)
Age groups		
Age25-30	0.034	(0.024)
Age30-35	0.051**	(0.026)
Age35-40	0.028	(0.043)
Age40-45	-0.010	(0.111)
Previous pregnancies		
Previous spontaneous abortions	-0.049**	(0.024)
Pregnancy-related variables		
Multiple Gestation	0.248	(0.155)
Abnormal presentation	0.114**	(0.059)
premature rupture of membrane	0.011	(0.028)
Placenta praevia	0.098	(0.114)
Maternal hypertension	-0.083	(0.226)
Post term delivery	-0.299***	(0.122)
Socio-economic status		
Income quintile 2	0.035	(0.027)
Income quintile 4	0.061***	(0.023)
Income quintile 5	0.025	(0.034)
Year dummies		
Year 2006	0.023	(0.042)
Year 2007	0.074*	(0.039)
Year 2008	0.065*	(0.038)
Year 2009	0.039	(0.038)
Intercept	-0.081*	(0.044)

N	1534
R^2 with physician effects	0.353
R^2 without physician effects	0.022

Significance levels : $*$: 10% $**$: 5% $***$: 1%

status significantly influences the likelihood of receiving a c-section. In fact, table 3.IV shows that for women of high socio-economic status (income quintile 4), the likelihood of a c-section increases by about 6 percentage points compared to women of low socio-economic status (income quintile 1, the reference group).

Finally, table 3.IV shows that the coefficients of determination (R^2) of the models with and without physician fixed effects are 0.353 and 0.022, respectively. More specifically, the results suggest that without physician fixed effects, the model explains only about 2% of the variance in c-section rates while including fixed effects substantially increases the percentage to nearly 35%. That is, the results suggest that physician factors are very important determinants of the c-section decision. In fact, physician factors account for about 94% of the R^2 of the model. Table 3.VII in the appendix D presents for each hospital, the R^2 of the models with and without physician fixed effects, as well as the percentage of the R^2 of the full model explained by physician fixed effects. It shows that in most of the hospitals (with a relatively high number of physicians), physicians fixed effects substantially increase the coefficient of determination.

The mean and the standard deviation of risk-adjusted cesarean rates are 0.137 and 0.277, respectively. That is, the deviation between a particular physician's risk-adjusted c-section rate and the hospital average is substantial and is about 28 percentage points, suggesting that about 28 percentage points of the variation across c-section rates in this specific hospital is due to physician practice styles. Table 3.V below reports the mean and standard deviation of risk-adjusted cesarean rates in each hospital.

I find that the standard deviation across physicians in most of the hospitals is substantial, between 0.4 and 27.7 percentage points, and in more than one-third of all hospitals, it is greater than 4 percentage points. In other words, the results suggest that, in these hospitals, more than 4 percentage points of the variation across c-section rates may be due to physician

Table 3. V – Risk-Adjusted Cesarean Rates

Total deliveries	Phys.	Mean	Std.Dev.	RSD
1,593	11	0.004	0.038	9.500
1,534	16	0.137	0.277	2.022
1,336	3	0.187	0.054	0.289
1,308	10	0.041	0.049	1.195
1,299	7	0.122	0.067	0.549
1,221	13	0.192	0.039	0.203
1,198	8	0.024	0.017	0.708
1,162	4	0.038	0.014	0.368
1,158	10	0.006	0.010	1.667
1,141	11	0.079	0.033	0.418
876	13	0.009	0.035	1.086
713	8	0.080	0.040	0.500
712	2	0.147	0.014	0.095
704	4	-0.023	0.029	1.261
679	6	0.076	0.093	1.224
665	7	-0.060	0.035	0.583
530	6	0.107	0.024	0.224
530	2	0.017	0.006	0.353
526	4	0.191	0.053	0.277
398	5	-0.027	0.023	0.852
365	3	0.043	0.043	1.000
354	3	0.171	0.014	0.082
352	5	-0.125	0.026	0.208
316	4	0.320	0.205	0.641
311	3	- 0.009	0.008	0.889
307	2	-0.005	0.008	1.600
303	4	0.149	0.040	0.268
274	4	0.137	0.025	0.182
265	2	-0.064	0.029	0.453
260	4	-0.031	0.034	1.097
216	3	-0.032	0.040	1.250
212	3	0.160	0.090	0.563
195	3	0.215	0.071	0.330
194	2	0.005	0.062	12.400
190	3	0.035	0.014	0.400
186	2	-0.114	0.061	0.535
182	3	0.127	0.084	0.661
176	2	0.083	0.048	0.578
171	2	0.091	0.026	0.286
150	2	0.001	0.008	8.000
143	2	-0.099	0.008	0.081
142	2	0.405	0.043	0.106
130	2	-0.077	0.004	0.052
122	2	0.003	0.027	9.000
117	2	0.108	0.122	1.130

practice styles. As a consequence, some patients are receiving c-sections simply because of their physician's type : women with identical observable characteristics but who patronize different obstetricians in the same hospital do not necessarily have the same likelihood of receiving a c-section. If the average risk-adjusted c-section in each hospital is considered as the appropriate treatment given the hospital's characteristics (such as hospital processes, norms and standards), then the results suggest that there will be welfare losses due to these variations. In order to reduce these welfare losses and to encourage the efficient provision of care, policies that tend to increase the standardization of procedures may be efficient.[57]

In order to compare the variations across hospitals in risk-adjusted c-section rates, I compute the relative standard deviation (RSD) (i.e., the absolute value of the coefficient of variation) in each hospital. In fact, because the means of risk-adjusted cesarean rates are drastically different from each other, the RSD is more appropriate for such comparison. The results suggest that the risk-adjusted cesarean rates vary substantially between hospitals.

Tables 3.VIII to XII (in appendix D) present the RSD of risk-adjusted cesarean rates by province. These results suggest that the risk-adjusted cesarean rates vary substantially within a same province as well as between provinces.

6 Conclusion

The question that motivated this paper is whether physicians differ in their propensity to perform a c-section. To this end, using data on deliveries in Canada between 2005 and 2009, I examine variations across physicians in c-section rates, controlling for observable patient characteristics. In order to disentangle the effect of physician treatment styles from hospital characteristics on these variations, the present paper focuses on within-hospital variations.

[57] It is important to note that, if women select their obstetrician according to their preference for the delivery method (c-section vs. vaginal birth), then a standardization may reduce their welfare. But since I restrict the analysis to a sample where women are less likely to select their obstetrician, this is not an issue in the present setting.

The paper also takes into account of the problem of patient selection or matching between physicians and patients by using institutional level data where patients are more likely to be randomly assigned to physicians and also by relying on a test of the randomness of patients to physicians.

The results show that physician factors are an important element in the c-section decision and that there is substantial variation across physicians in c-section rates within the same hospital. More specifically, the results show that the standard deviation of risk-adjusted cesarean rates is between 0.4 and 27.7 percentage points. That is, the results suggest that physicians differ substantially in their propensity to perform a c-section and the delivery method (c-section vs. vaginal birth) substantially depends on obstetrician effect. Finally, the results suggest that the risk-adjusted cesarean rates vary substantially within the same province as well as between provinces.

The results point to an interesting avenue which may encourage the efficient provision of care. In fact, since the deviation from the average risk-adjusted c-section in each hospital may represent a welfare loss, then policies that encourage hospital norms and standards can be efficient.

Although the results suggest that, in the same hospital there is an important variation across physicians in c-section rates, it may be interesting to analyze the sources of these variations. That is, one may decompose the total variation into components that are attributed to each physician characteristics (such as age, experience, gender and medical training). However, it is impossible with the DAD, because physician identifiers can not be merge with the Scott's Medical Database (SMDB) which contains information on physician characteristics.

7 References

[1] Baldwin Laura-Mae, Hart L. Gary, Lloyd Michael, Fordyce Meredith and Rosenblat A. Roger (1995) : 'Defensive medicine and obstetrics' , *Journal of the American Medical Association*, 274, 1606- 1660.

[2] Burns R. Lawton, Geller E. Stacie and Wholey R. Douglas (1995) : 'The effect of physician factors on the cesarean section decision', *Medical care*, 33 (4), 365-382.

[3] Bynum Julie, Song Yunjie and Fisher Elliot, (2010) : 'Variation in prostate-specific antigen screening in men aged 80 and older in fee-for-service medicare', *Journal of the American Geriatrics Society*, 58(4), 674-680.

[4] Dubay Lisa, Kaestner Robert and Waidmann Timothy (1999) : 'The impact of malpractice fears on cesarean section rates', *Journal of Health Economics*, 18 (4), 491-522.

[5] Epstein J. Andrew and Nicholson Sean (2009) : 'The formation and evolution of physician treatment styles : an application to cesarean sections', *Journal of Health Economics,* 28, 1126-1140.

[6] Epstein J. Andrew, Ketcham D. Jonathan and Nicholson Sean (2010) : 'Specialisation and matching in professional services firms', *RAND Journal of Economics,* 41(4), 811-834.

[7] Fisher S. Elliot, Bynum P. Julie, Skinner S. Jonathan (2009) : 'Slowing the growth of health care costs- lessons from regional variation', *New England Journal of Medicine,* 360 (9), 849-852.

[8] Grant Darren (2005) : 'Information and sorting in the market for obstetrical services', *Health Economics*, 14, 703-719.

[9] Grant Darren (2009) : 'Physician financial incentives and cesarean delivery : new conclusions from the healthcare cost and utilization project', *Journal of Health Eco-*

nomics, 28 (1), 99-123

[10] Gregory D. Kimberly, Korst M. Lisa, Gornbein A. Jeffrey and Platt D. Laurence Darren (2002) : 'Using administrative data to identify indications for elective primary cesarean delivery', *Health Services Research*, 37 (5), 1387-1401.

[11] Groenewegen P. Peter. and Westert P. Gert (2004) : 'Is there a time trend in medical practice variations ? : a review of the literature and an critical analysis of theoretical approaches', *Journal of Health Economics*, 28 (1), 99-123

[12] Gruber Jonathan and Owings Maria (1996) : 'Physician financial incentives and cesarean section delivery', *Rand Journal of Economics,* 27 (1), 99-123.

[13] Gruber Jonathan, Kim John and Mayzlin Dina (1999) : 'Physician fees and procedure intensity : the case of cesarean delivery', *Journal of Health Economics*, 18 (4), 473-490.

[14] Grytten Jostein and Sorensen Rune (2003) : 'Practice variation and physician-specific effects', *Journal of Health Economics*, 22 (3), 403-418.

[15] Hanan L. Edward, Wu Chuntao and Chassin R. Mark (2006) : 'Differences in per capita rates of revascularization and in choice of revascularization procedure for eleven states', BMC *Health Services Research*, 6 (35),1-8

[16] Henry A. Olivia, Gregory D. Kimberly, Hobel J. Calvin and Platt D. Laurence Darren (1995) : 'Using ICD-9 codes to identify indications for primary and repeat cesarean sections : agreement with clinical records', *American Journal of Public Health*, 85 (8), 1143-1445.

[17] Horwitz M. Sarah, Mitler K. Lloyd and Rizzo A. John (2000) : 'Physician gender and cesarean sections', *Journal of Clinical Epidemiology*, 53, 1030-1035.

[18] Localio A. Russell, Lawthers G. Ann, Bengtson M. Joan, Hebert E. Liesi, Weaver L. Susan, Brennan A. Troyen and Landis J. Richard (1993) : 'Relationship between

malpractice claims and cesarean delivery', *JAMA*, 269, 366-373.

[19] McMahon F. Laurence, Wolfe A. Robert, Haung Saling, Tedeschi Philip, Maning Willard and Edlund J. Mark (1999) : 'Racial and gender variation in use of diagnostic colonic procedures in the Michigan medicare population', *Medical care*, 37 (7), 712-717.

[20] Neyman J. and Scott L. Elizabeth (1948) : 'Consistent estimates based on partially consistent observations', *Econometrica*, 16 (1), 1-32.

[21] Phelps E. Charles (2000) : 'Information diffusion and best practice adoption', in *Handbook of Health Economic*s.

[22] Phelps E. Charles and Parente T. Stephen (1990) : 'Priority setting for medical technology and medical practice assessment', *Medical Care* 28 (8), 703-723.

[23] Sloan A. Frank, Entman S. Stephen, Reilly A. Bridget, Glass A. Cheryl, Hickson B. Gerald and Zhang H. Harold (1997) : 'Tort liability and obstetricians' care levels', *International Review of Law and Economics*, 17, 245-260.

[24] Tu V. Jack, Ko T. Dennis, Guo Helen, Richards A. Janice, Walton Nancy, Natarajan K. Madlu, Wijeysundera C. Harinda, So Derek, Latter A. David, Feindel M. Christopher, Kingsbury Kori , Cohen A. Eric (2012) : 'Determinants of variations in coronary revascularization practices',*Canadian Medical Association Journal*, 184 (2), 179-186.

[25] Tussing A. Dale and Wojtowycz Martha (1992) : 'The cesarean decision in New York state, 1986 : economic and noneconomic aspects', *Medical care*, 30 (6), 529-540.

[26] Wennberg John and Gittelsohn Alan (1973) : 'Small area variations in health care delivery', *Science*, 182, 1102-1108.

[27] Wennberg John, Fisher S. Elliot and Skinner S. Jonathan (2002) : 'Geography and the debate over medicare reform', *Health Affairs*, Web exclusive, W96-W114.

[28] World Health Organization : 'Appropriate technology for birth', *Lancet* 1985 ; 2 : 436-7.

Table 3. VI – Number of Deliveries, Physicians and C-Section Rates by Hospital

Total deliveries	**Phys.**	**c-section rate**
1,593	11	0.102
1,534	16	0.300
1,336	3	0.103
1,308	10	0.190
1,299	7	0.106
1,221	13	0.124
1,198	8	0.019
1,162	4	0.050
1,158	10	0.022
1,141	11	0.232
876	13	0.146
713	8	0.158
712	2	0.094
704	4	0.111
679	6	0.125
665	7	0.168
530	6	0.083
530	2	0.038
526	4	0.093
398	5	0.116
365	3	0.096
354	3	0.037
352	5	0.145
316	4	0.152
311	3	0.029
307	2	0.065
303	4	0.089
274	4	0.161
265	2	0.034
260	4	0.165
216	3	0.019
212	3	0.090
195	3	0.144
194	2	0.165
190	3	0.274
186	2	0.146
182	3	0.071
176	2	0.170
171	2	0.053
150	2	0.027
143	2	0.168
142	2	0.317
130	2	0.031
122	2	0.115
117	2	0.197

Table 3. VII – R Squared with and without Physician Fixed Effects

Total deliveries	**Phys.**	**R^2 without**	**R^2 with**	**Perc.**
1,593	11	0.081	0.095	0.147
1,534	16	0.022	0.353	0.938
1,336	3	0.054	0.078	0.309
1,308	10	0.256	0.275	0.070
1,299	7	0.110	0.132	0.167
1,221	13	0.203	0.214	0.051
1,198	8	0.053	0.063	0.159
1,162	4	0.278	0.282	0.014
1,158	10	0.299	0.306	0.023
1,141	11	0.398	0.404	0.015
876	13	0.332	0.339	0.021
713	8	0.454	0.460	0.013
712	2	0.069	0.070	0.014
704	4	0.521	0.526	0.010
679	6	0.080	0.138	0.420
665	7	0.621	0.628	0.011
530	6	0.434	0.445	0.025
530	2	0.226	0.228	0.009
526	4	0.083	0.099	0.162
398	5	0.274	0.280	0.021
365	3	0.350	0.365	0.041
354	3	0.135	0.142	0.049
352	5	0.429	0.440	0.025
316	4	0.099	0.327	0.697
311	3	0.508	0.511	0.004
307	2	0.186	0.187	0.005
303	4	0.382	0.402	0.005
274	4	0.279	0.280	0.004
265	2	0.499	0.504	0.010
260	4	0.431	0.435	0.009
216	3	0.430	0.446	0.036
212	3	0.271	0.331	0.181
195	3	0.183	0.205	0.107
194	2	0.243	0.254	0.043
190	3	0.507	0.510	0.006
186	2	0.436	0.438	0.005
182	3	0.115	0.174	0.339
176	2	0.676	0.681	0.007
171	2	0.373	0.376	0.008
150	2	0.097	0.098	0.010
143	2	0.780	0.781	0.001
142	2	0.268	0.272	0.015
130	2	0.509	0.510	0.002
122	2	0.632	0.635	0.005
117	2	0.423	0.462	0.084

Table 3. VIII – Variation in Risk-Adjusted Cesarean Rates-Ontario

Total deliveries	**Physicians**	**RSD**
1,336	3	0.289
1,221	13	0.203
1,198	8	0.708
1,141	11	0.418
704	4	1.261
665	7	0.583
530	6	0.224
526	4	0.277
365	3	1.000
352	5	0.208
311	3	0.889
303	4	0.268
260	4	1.097
195	3	0.330
176	2	0.578
142	2	0.106
122	2	9.000

Table 3. IX – Variation in Risk-Adjusted Cesarean Rates-Alberta

Total deliveries	**Physicians**	**RSD**
1,593	11	9.500
1,534	16	2.022
1,308	10	1.195
1,299	7	0.549
679	6	1.224
398	5	0.852
354	3	0.082
316	4	0.641
307	2	1.600
274	4	0.182
194	2	12.400
182	3	0.661
171	2	0.286
150	2	8.000
130	2	0.052
117	2	1.130

Table 3. X – Variation in Risk-Adjusted Cesarean Rates-Manitoba

Total deliveries	**Physicians**	**RSD**
1,162	4	0.368
1,158	10	1.667
713	8	0.500
530	2	0.353
216	3	1.250

Table 3. XI – Variation in Risk-Adjusted Cesarean Rates-Saskatchewan

Total deliveries	**Physicians**	**RSD**
712	2	0.095
265	2	0.453
212	3	0.563

Table 3. XII – Variation in Risk-Adjusted Cesarean Rates-Nova Scotia

Total deliveries	**Physicians**	**RSD**
876	13	1.086
186	2	0.535

Conclusion générale

Dans cette thèse qui comporte trois essais, nous avons analysé à travers des modèles théoriques et empirique différentes questions reliées au comportement des médecins et différents mécanismes pouvant permettre d'atteindre l'efficacité dans le secteur des services de soins de santé.

Dans le premier essai, nous avons examiné dans quelle mesure la menace d'un patient de changer de médecin en cas de soins de santé inadéquats peut encourager l'efficacité dans un environnement où les médecins sont rémunérés par capitation. Compte tenu de l'ignorance des patients et de la nature de l'information qu'ils reçoivent sur l'éventuel futur médecin, les résultats indiquent qu'ils commettent des erreurs de type I et II. Les résultats indiquent également que la menace de changer de médecin n'incite pas forcément le médecin à traiter son patient de façon appropriée. Ainsi, certains médecins peuvent fournir moins de soins que ce qui est approprié et d'autres peuvent même pratiquer une forme de médecine défensive en fournissant trop de traitement.

Les résultats suggèrent des pistes intéressantes pouvant encourager une offre et une consommation efficaces des soins de santé. En effet, les résultats suggèrent que les mesures visant à réduire l'ignorance des patients par rapport à leur maladie doivent être encouragées. Les résultats suggèrent également que les mesures visant à fournir de l'information sur la qualité des médecins pourraient augmenter le bien-être des patients. Finalement, toute action visant à réduire les coûts liés à la recherche d'un nouveau médecin devrait être encouragée.

Dans le deuxième essai, nous avons développé un modèle qui prend en compte l'interdépendance entre les systèmes de rémunération des médecins et les mécanismes de responsabilité pour faute médicale. Les résultats indiquent que le fait que la cour de justice puisse se tromper en sanctionnant de façon injuste un médecin qui a fourni les soins appropriés incite les patients à poursuivre leurs médecins même s'ils savent qu'ils ont reçu le traitement approprié. En conséquence, les médecins répondent à cette menace en pratiquant la médecine défensive. Les résultats suggèrent donc que les mesures visant à réduire les décisions inappropriées de la cour pourraient permettre d'atteindre l'efficacité.

Les résultats montrent également que, à cause des frais liés à la poursuite des médecins, les patients qui sont averses au risque ont beaucoup plus de réticence à poursuivre leurs médecins. Ces patients sont prêts à transférer les responsabilités de poursuite à une firme d'avocats en contrepartie d'un certain pourcentage du montant de la compensation anticipée. Par conséquent, des politiques pouvant permettre aux patients de s'assurer pour les poursuites pour faute médicale pourraient permettre d'augmenter leur bien-être.

Dans le troisième essai, nous avons examiné de façon empirique dans quelle mesure les variations dans les taux de césarienne à l'intérieur des hôpitaux sont expliquées par les facteurs relatifs aux médecins. De façon générale, les résultats montrent que les facteurs propres aux médecins expliquent une partie importante de la variabilité totale dans les taux de césarienne. Ces résultats suggèrent donc que les caractéristiques des médecins sont décisives dans le choix de la méthode d'accouchement. Par conséquent, toute politique publique visant à réduire le taux de césarienne à un niveau qui soit plus en lien avec le niveau recommandé par l'OMS devrait forcément tenir compte des caractéristiques des médecins.

De plus, en contrôlant pour les conditions médicales ainsi que les autres caractéristiques observables des patients, les résultats indiquent que la probabilité d'accoucher par césarienne pour un patient donné dépend fortement du type de son médecin. En d'autres termes, les résultats suggèrent que, dans un même hôpital, deux patients identiques ayant les mêmes caractéristiques ont des chances différentes d'accoucher par césarienne s'ils sont traités par des médecins différents. En effet, dans la plupart des hôpitaux, nos résultats montrent d'importantes variations dans les taux de césarienne ajustés par rapport au risque. Les résultats suggèrent donc que les politiques qui encouragent une standardisation des procédures à l'intérieur des hôpitaux pourraient augmenter le bien-être social.