Laboratory Kickback to Doctor, Consumer Awareness, and Welfare Implications*

Yuk-fai Fong[†], Xiaoxuan Meng[‡] and Lin Zhao[§]

January 29, 2020

Abstract

We consider a model in which the doctor conducts imperfect and unverifiable diagnosis and clinical laboratories provide tests to enhance the precision of diagnosis at the same time makes diagnosis verifiable. Kickback from laboratory motivates the doctor to over-prescribe laboratory tests. Doctor may cheat on both severity and the type of problem. As a result, wrong and unnecessary tests are conducted. Part of patients being aware of kicback results in less kickback while not necessarily mitigate the under-prescription problem. Banning kickback can hurt social welfare when under-prescription is a severe problem in the no kickback senario.

1 Introduction

Credence goods literature has broadly studied the doctor's incentive to misreport patients' problems to overcharge and/or over-treat them. The existence of laboratory test can enhance the precision of the doctor's diagnosis and makes the diagnosis verifiable. Independent laboratories are prevalent and increasingly involved in medical service. It has been reported that approximately 80% of physicians' diagnoses are a result of laboratory

*The authors would like to acknowledge the Hong Kong RGC Grant (Project Number 16500018) for its financial support.

†The University of Hong Kong, email: yukfaims@hku.hk

‡The University of Hong Kong, email: xxmeng@hku.hk

§University of Hong Kong, email: linzhao@hku.hk

tests[1]. The laboratory market is highly competitive. *Laboratory Corporation of America*, *Quest Diagnostics Inc. Alere, Inc., Bio-Reference Laboratories, DaVita Healthcare Partners, Inc.* and *Genomic Health, Inc.* are all leading clinical laboratory companies providing common routine test as well as specialty test. It is common for medical laboratories to pay doctors kickbacks to incentivize them to direct patients their way. Kickback from medical laboratories to doctors becomes a big concern for the society and can cause a serious problem. According to Wall Street Journal (2014), "A fast-growing Virginia laboratory has collected hundreds of millions of dollars from Medicare while using a strategy that is now under regulatory scrutiny: It paid doctors who sent it patients' blood for testing." Department of Justice in U.S. announced that "Quest Diagnostics Inc. has agreed to pay $6 million to resolve a lawsuit by the United States alleging that Berkeley HeartLab Inc., of Alameda, California, violated the False Claims Act by paying kickbacks to physicians and patients to induce the use of Berkeley for blood testing services and by charging for medically unnecessary tests."[2] Such kickbacks are also not uncommon outside of the United States. For instance, according to Wong (2010), kickbacks from laboratories to doctors are alleged to happen in Hong Kong.

Despite the ubiquity of clinical laboratories and kickbacks, little is known about the potential conflict of interest: when does doctor have incentives to over-prescribe or under-prescribe laboratory tests, the purpose of which is predominately to provide a more accurate diagnosis? Only a small branch of literature examines referral incentives. Pauly(1979) discussed the kickback paid for patients referrals from one physician to another. Inderst and Ottaviani (2012) developed an elegant setting in which sellers have to rely on an information intermediary to recommend their products to consumers. However, their model does not capture some important aspects of referral of laboratory tests. This is because both the doctor and the medical laboratories are sellers of services as well as information intermediaries. They're both sellers because the doctor sells treatments and the med-

[1] For more information, see https://blog.marketresearch.com/12-leading-companies-in-clinical-laboratory-services

[2] For more information, see https://www.justice.gov/opa/pr/blood-testing-laboratory-pay-6-million-settle-allegations-kickbacks-and-unnecessary-testing

ical laboratories sell lab tests. They are both information intermediaries because only the doctor knows whether it is cost effective to refer a patient to the laboratory and the medical laboratory improves the accuracy of the doctor's diagnosis and makes the diagnosis verifiable. To capture the feature of relative information advantage between lab and doctor, we build a model based on the standard credence goods model (see Pitchik and Schotter(1987), Wolinsky(1993), Fong(2005)), in which two treatments are available to fix patients' problem which may be either of two possible levels of severity. We further extend the model to consider another dimension of information of the problem: type. After patient's visit, doctor reports the diagnosis containing two pieces of information: severity and type, and then recommends either treatment directly or a laboratory test according to the diagnosis. Diagnosis for the type is perfect while the severity is not. Lab can only get a reliable result based on the right information of type. Hence doctor's diagnosis is necessary for the lab test. We show that the interaction between doctor and patients in a market with laboratories can be qualitatively different from results in standard credence goods model without laboratories. Monitoring by rejection (See Fong(2005), Dulleck and Kerschbamer(2006)) is no longer necessary to result in market ineffeciency. Truth-telling always happens when there is no commissions from labs to the doctor. In this paper, we put our major attention on how kickback influences the referral incentive for the doctor.

We further assume that heterogeneous patients suffer differently from serious problem if they are left untreated. High-loss patients suffer more while low-loss patients suffer less. In a truth-telling equilibrium, under-provision of lab test can happen when wrong diagnoses occur less often. As the precision of the doctor's diagnoses improved, patients are willing to pay more when they get an inaccurate diagnoses. The expected willingness-to-pay for those high-loss patients are much higher than those low-loss patients. A profit maximized doctor cares more about high-loss ones. As a result, some low-loss patients are not treated. Kickback shifts doctor's interest in referring patients and results in an unnecessary testing and wrong testing problem. Lab tests are over-provided to some patients, especially when patients are unaware of kickback. Doctor tries to exploit unawareness and induces more patients to conduct test regardless of diagnoses. It is interesting to see that under-provision and over-provision can happen at the same time. When probability of serious problem is high, doctor choose to induce more test among high loss patients while no test among low loss patients. Because more profit earn from high-loss patients outweighs the loss from uncovered of the low-loss ones.

Patients awareness can efficiently reduce kickback. However, the impact of awareness on the social welfare highly depends a lot of parameters such as precision of diagnoses and cost of lab test. When cost of lab test is low, being aware of kickback will mitigate under-provision of lab tests and result in more serious problems being treated in the equilibrium. In contrast, when cost of lab test is high, being aware of kickback leads to less patients being treated in the equilibrium. As patients become aware of kickback, their beliefs about the actual problem they got become less biased, so that there is less room for the doctor to exploit through kickback.

We also discuss some other policies such as banning kickback and price cap on the prices. Banning kickback can hurt social welfare when the under-provison of lab test is severe in the no kickback case. A small limitation on the price of treatment helps improve the under-provision problem while limitation of treatment prices may hurt social welfare.

2 Literature Review

Since the seminal work of Darby and Karni (1973), feature of credence goods has being well explored in theoretical works. Wolinsky(1993) studies the role of search and reputation in regulating the incentive of expert. Emons(1997) presents some nonfraudulent equilibrium when the treatment are verifiable. Fong(2005) also shows some no cheating equilibrium with consumer's rejection to the serious treatment, which results in market inefficiency. Liu(2011) adds some behavior types of expert to explain the market inefficiency. A comprehensive discussion on credence goods can be found in Dulleck and Kerschbamer(2006). Our model is built on the standard model of credence goods with liability assumption (see, for example, Wolinsky(1993), Fong(2005)). Papers we mentioned above focus on the problem of expert's fraud on the suitability of treatment due to the asymmetric information problem. This is not our focus as we introduce a lab who can provide extra information to eliminate the uncertain about the suitability of treatment. We mainly study the problem caused by asymmetric information about the necessity of lab test.

In our paper, doctor acts as an information intermediary and the demand of lab test highly depends on the diagnostic information as well as recommendation from the doctor. Most existing developments on the theory of information intermediaries do not appear to be directly applicable to the medical care service market we study in this paper. A large stream of works focus on pure advisory role of the information intermediatries. Armstrong (2011) considers a senario that firms pay commissions to the advisor in order to get a postion of prominent. Uninformed consumers rely on the adviors' information and are directed to more expensive sellers. When sellers differ in the quality or the cost of the product they supply, paying for the prominent improve market performance as it improves the effectiveness of consumer search. The firm with highest quality is willing to pay the most to become prominent.

In Inderst and Ottaviani (2012a), firms use kickback to steer the advice from the intermediary. More cost-efficient firm is able to give more kickback, resulting in a larger market share. In our paper, due to the informational role of the lab and the un-verifiability of the treatment, doctor can make use some unobservable information and earn more profit by charging different kickbacks from different labs, even when the two labs are equally cost-efficient. Inderst and Ottaviani (2012AER) also discussed the impact of disclosing commission to consumers. When consumer observes the commissions by each firm, they can directly infer the recommendation rule of the intermediary. The price that consumer is willing to pay depends on this recommendation rule. Any change in the observable commission will be reflected in price immediately. Disclosure of commissions does not affect advice when firms are equally cost-effcient. However, in our paper, the patients' decision is only about choosing between two products. A correct belief about kickback and recommendation rule leave no room for the doctor to exploit the patient though kickback. Given the same prices, more patients rejects the recommendation for the lab test. As both lab and doctor are sellers of product, the demand of lab test will affect the demand for the treatment offered by the doctor. Then no kickback is profitable for the doctor if patients can observe the kickback. Recommendation rule will change dramatically due to the disclosure of kickback.

Although there are few other works consider a multi-role information intermediaries, these two roles are mutually substitutive. Pauly(1979) discussed the referral incentive for physicians in medical care market. In his model, patients first visit a generalist. And the generalist then decides to treat the patients himself or refer the patient to a specialist who has a lower marginal cost then him. Kickback from specialists changes the number of referrals since generalist finds that treating the patient himself is sometimes less profitable than referring them to a specialist and getting the kickback. Kickback improves the patient welfare since if kickback is permitted, the least-cost combination of the treating plan will be adopted. The incentive of referral in our paper is quite different from Pauly(1979). Doctor and labs are providing partially complementary products instead of substitutes as in Pauly(1979). Kickback increases the referral patients even when it is not cost efficient to do so.

Condorelli et.al(2018) studies information intermediaries in an aution market. Sellers have no direct access to buyers. These intermediaries can either bid the object theseselves and reselling it to his buyers or refer his buyers to the seller. When seller has full bargaining power, in the equilibrium, with any positive commissions from seller to intermediaries, the intermediaries only adopt referral mode. As a result, mis-allocation distortion caused by resale is eliminated. Both industrial profit and consumer surplus are higher than in the case that intermediaries resells the object.

There are a few strands in the development of IO models with naivety of consumers (Spiegler (2011)). Some naiveties are about the dynamically inconsistent preferences (see, for example, Strotz (1956), Peleg and Yaari (1973)). Naive consumers fail to correctly estimate the likelihood of future taste changes. Another type of naivety is related to incapability of the bounded rational consumers dealing with market complexity. For example, Inderst and Ottaviani (2012b) consider the situation that some consumers naively believe that expert's advice is unbiased. When naive consumers observe high price, they can not related it to higher commission. As a result, they under-estimate the probability of receiving the recommendation of this high commission and more expensive product. Our study falls into the latter category of naivety . In our model, naivety (unawareness) about the kickback leads to an over-estimation of the probability of the problem that they are willing to pay more when they are facing a recommendation of lab test. Naivety of the patients are exploited by charging a low price for the problem being over-estimated while a high commission from labs.

3 Baseline Model

3.1 Doctor, problem and lab test

There is a continuum of patients with measure one. Each patient (henceforth she) has a problem, the nature of which is captured by two dimensions. First, *type* of the problem, $T \in \{A, B\}$. Each type of problem happens with equal probability. The other dimension is the *severity* of the problem, denoted by i, which is either minor (m) or serious (s). Regardless of the type (T) of the problem, if a problem with severity $i \in \{m, s\}$ is left untreated, a patient has to bear a loss of l_i, where $0 < l_m < l_s$.

Although patients may be infected with the same disease, subjective feeling of each patient differs. Assume that the loss from a serious problem follows a uniform distribution $l_s \sim F(\cdot)$ with support $l_s \in [\underline{l}_s, \bar{l}_s]$. l_s is private information of each patient while the distribution is common knowledge among all players. A minor problem which is left untreated will lead to same loss for every patient.

There is a monopolist doctor (henceforth he) who provides diagnosis and treatment services. Diagnosis is costless but imperfect in assessing the severity of the problem. After the doctor performs a diagnosis, he learns the exact value of T but only obtain an imperfect signal about severity, $d \in \{m, s\}$: when the problem is serious, the doctor observes $d = s$ with probability one, i.e., $\Pr(d = s|i = s) = 1$; however, when the problem is minor, $\Pr(d = m|i = m) = \theta \in (0.5, 1)$.

There are two types of treatments, $T_t \in \{A, B\}$, corresponding to the two types of problem. Within each type of problem, there are two levels of treatments: minor treatment and serious treatment. If $T_t \neq T$, then the problem cannot be treated. Conditional on

$T_t = T$, the minor treatment can only treat the minor problem while the serious treatment can treat both problems. Doctor can only set a unique price for each treatment. Screening is not allowed here. It is common knowledge that for any type of problem, with probability α the patient's problem is serious. The costs of these treatments are respectively denoted by $c_m \in (0, l_m)$ and $c_s \in (0, l_s)$, which are common across types of treatments.

Each laboratory offers one type of test, $T_L \in \{A, B\}$, in a competitive market. We denote them as lab A and lab B. After conducting a test, the result shows whether the patient gets a serious problem for the respective type of problem. Laboratories do not have the ability to diagnose the type of problem T. We assume that if a lab does not know the type of a patient's problem and provides a wrong type of test, i.e., $T_L \neq T$, the test is completely uninformative. In details, it always shows the patient does not get a serious problem of type T_L. On the other hand, if $T_L = T$, a lab can perfectly diagnose the severity of a patient's problem. Note that the doctor's diagnosis and the lab test are complementary. Combined with information from a truth-telling doctor, the reuslt from lab test is enough to show an accurate diagnosis for the paitent's problem. Given our assumptions, there is no value for a patient to go to a laboratory before seeing an doctor. Once the doctor identifies the problem's type, the lab test both enhances the precision of the doctor's diagnosis and makes the diagnosis verifiable. We say that a test is necessary when doctor's diagnosis is imperfect. The cost of performing a test is $c_L > 0$.

Although treatment provided by the doctor is credence good. As long as there are labs who can provide more information about the problem, asymmetric information is resolved at most of time after the conduction of lab test[3].

We restrict our attention to situations in which the following commonly assumed conditions are satisfied:

$$\begin{gathered} 0 < c_m < l_m, \quad 0 < c_s < \underline{l}_s, \\ \alpha \bar{l}_s + (1 - \alpha)\, l_m < c_s. \end{gathered} \tag{R}$$

One immediate implication of (R) is $0 < c_m < l_m < c_s < \underline{l}_s < \bar{l}_s$. The first line of (R) states that the doctor has cost-effective technologies to treat both problems and the

[3] We will show in the later discussion that sometimes a wrong lab test will be conducted in the equilibrium. Asymmetric information still exist after this wrong test.

second line rules out uninteresting cases.

In our model, when doctor's diagnosis of severity is serious, with probability of $\frac{(1-\alpha)(1-\theta)}{\alpha+(1-\alpha)(1-\theta)}$, the problem is actually minor. Without a lab test, to ensure this patient is treated, a serious treatment must be conducted. However, when a lab test can give a perfect diagnosis of the problem, the expected cost of treat the problem is becomes $c_L + \frac{\alpha c_s+(1-\alpha)(1-\theta)c_m}{\alpha+(1-\alpha)(1-\theta)}$ instead of c_s. We further assume that it is socially efficient to conduct laboratory test when patients' problems are not perfectly diagnosed. That is

$$c_s > c_L + \frac{\alpha c_s + (1-\alpha)(1-\theta)c_m}{\alpha + (1-\alpha)(1-\theta)}$$

It can be rewrite as:

$$\frac{(1-\alpha)(1-\theta)}{\alpha + (1-\alpha)(1-\theta)}(c_s - c_m) > c_L \tag{1}$$

We put some further restriction on the cost to simplify analysis: $\alpha(\underline{l}_s - c_s) > c_L$.

3.2 Contract between labs and doctor

Suppose laboratories are independently operated. There are competitive segments of the market for two types of lab tests. Within each segmentation, labs provide exactly same diagnostic testing service at same cost. Bertrand competition results in zero profit. It is equivalent to assume that the doctor proposes a "take-it-or-leave" offer to the lab.

Suppose the doctor cooperates with two labs (Lab A and Lab B) who provide tests for problem A and B respectively. The contract between doctor and labs are a contingent payment conditional on each referal of patient. As shown in the Figure 1, as long as doctor and lab make agreement on the kickback, whenever the a patient is referred to a lab, no matter whether it is a necessary one, kickback will be paid to doctor.

We should also emphase here that the contract between doctor and labs are informal and unenforcable. Kickback should be self-enforcing.

3.3 Timeline

The game proceeds as follows. At the beginning of the game, the doctor and labs agree on the kickback for referral of lab test (if there is any). The doctor announces prices for

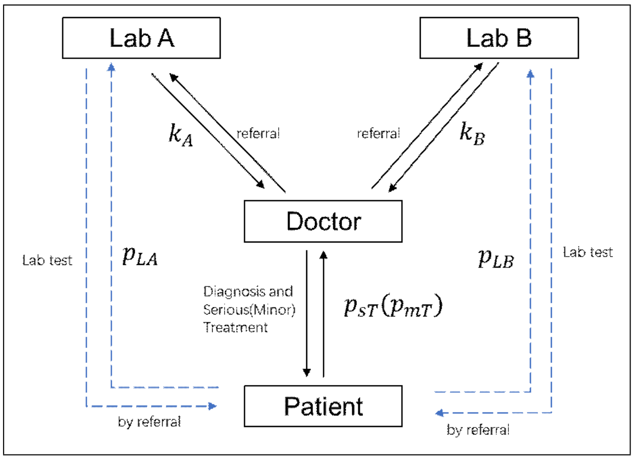

Figure 1: This scheme illustrates actions and payment aong the labs, doctor and patients.

the treatments: $\{p_{mT}, p_{sT}\}_{T\in\{A,B\}}$. Labs annouces price of lab test, p_{LT}, $T \in \{A, B\}$. Upon the observation of all the fees, patients with problems decide whether to visit the doctor. Following Fong (2005), we call the rest of the game from the moment the patient visits the doctor the *recommendation subgame*, but there are differences in details due to imperfect diagnosis and availability of laboratory test. Upon seeing the patient, the doctor first learns the type of problem T and then gets a signal d for severity. Then he (i) recommends a treatment at price $p \in \{p_{mT}, p_{sT}\}_{T\in\{A,B\}}$ (ii) refers the patient to receive a laboratory test at some lab, or (iii) refuses to provide any treatment[4]. Denote these recommedations as t. If the doctor recommends a treatment or refers the patient to a laboratory, then the patient decides whether to accept or reject the offer. If the patient accepts a treatment recommendation, the doctor must provide a treatment that ensures treating the problem at the price he has quoted. We assume that the existence of patients' problems is verifiable; however, patients do not know which of the two treatments has been actually provided as long as the problem is treated. As stated earlier, this treatment, however, may not have to be the one he has claimed to provide. If the doctor refers the patient to a laboratory and the patient accepts the referral, then the patient has to pay p_{LT}.

[4]Following Wolinsky (1993) and Fong (2005), we assume that at this point of the game the expert offers to treat the customer's problem at either p_{mT} or p_{sT} only if doing so is profitable in expectation.

3.4 Patient rationality

Throughout the whole paper, we distinguish between two types of patients: unaware and aware. Aware patients notice the posibility that a doctor may receive kickback from labs and they perfectly understand the incentives arising from the commissions. To be specific, we assume that the contract between lab and doctor is not disclosed to patients unless some disclosure policies are adopted. Aware patients, nevertherless, form rational beliefs.

Unaware patients naively believe that there is no kickback between doctor and lab. As a result, incentives for the doctor to recommend unnecessary and wrong test are ignored by unaware patients. In the equilibrium, unaware patients may hold biased beliefs about kickback and doctor's recommendation strategy regarding test.

Bounded rationality only appears in aspect of kickback. All patients rationally anticipate the doctor's incentives to cheat on the treatment that actually be done because of the unverifiability. For example, doctor may want to charge a high price when actually conduct a minor treatment. We assume that both unaware and aware patients hold correct belief which is consistent with the doctor's strategy about recommendation of serious treatment. More specifically, if direct recommendation of a serious treatment is a part of on-equilibrium path strategy of doctor, then both aware and unaware patients believe that the problem is actually serious if

$$\bar{\gamma}_{sT}(p_{sT} - c_m) \leq p_{mT} - c_m.^5$$

If direct recommendation of a serious treatment only appears off the equilibrium path, then it is without generality to assume that patients believes that the problem is actually minor. This belief is consisitent with doctor's equilibirum strategy.

[5] $\bar{\gamma}_s$ is the acceptence rate of serious treatment. It is defined as following:

$$\bar{\gamma}_{sT} = \int_{\hat{l}_s}^{\bar{l}_s} \gamma_{sT}(l_s) dF(l_s) = \frac{\bar{l}_s - \hat{l}_{sT}}{\bar{l}_s - \underline{l}_s}$$

$$\hat{l}_{sT} = \min\{l_s \in [\underline{l}_s, \bar{l}_s] : p_s \leq \alpha_s \hat{l}_s + (1 - \alpha_s) l_m\}$$

α_s is the equilibrium belief which equal to the expected probability that the problem is serious conditional on a serious treatment is recommended.

3.5 Solution concept

The appropriate equilibrium concept is Perfect Bayesian Equilibrium (PBE) consisting of labs' strategies, doctor's strategy, patients' strategy and patients' beliefs. The equilibrium strategy of the doctor specifies a price list $\{p_{sT}, p_{mT}, k_T\}_{T=A,B}$ and, the probabilities $\{\beta_{dT}^{T'}, \rho_{dT}^{T'}, 1 - \sum_{T'} \beta_{dT}^{T'} - \sum_{T'} \rho_{dT}^{T'}\}_{d\in\{m,s\},T,T'\in\{A,B\}}$[6]. The equilibrium strategy of labs is a the price of lab test p_{LT}, for $T = A, B$. The equilibrium strategy of the patients specify a list of decition rule $\{\gamma_t(l_m, l_s), 1 - \gamma_t(l_m, l_s)\}_{t\in\{m,s,L\}}$ for $p_t \in \{p_{sT}, p_{mT}, p_{LT}\}_{T=A,B}$. We restrict attention to the optimal equilibrium, or the PBE which gives the doctor highest profit. Moreover, for ease of exposition, we adopt the tie-breaking rule that the doctor provides service when he earns the same profit whether he provides it or not.

4 Benchmark case without kickback

Since we are interested in analyzing some generic features of the market for lab test, we first consider an environment that kickback is prohibited. As there is no kickback, we do not specify whether the patients are aware of kickback. Absent the kickback, all patients should hold corret beliefs in the equilibrium.

Before charaterizing the optimal equilibrium for the doctor, we first want to figure out some essential features of the optimal equilibrium.

Lemma 1 *(No direct serious treament recommendation) In the optimal equilibrium without kickback, doctor never recommends a serious treatment directly on the equilibrium path.*

Since treatments are unverifiable as long as the problem has been treated. If a serious treatment recommendation appears on the equilibrium path, price as well as the acceptence rate of the serious treament will be largely restricted by following incentive compatibility condition:

[6] $\beta_{dT}^{T'}$ ($\rho_{dT}^{T'}$)denotes probability that doctor recommends a serious treatment (lab test) of type T' when the diagnosis is d and the type of problem is T.

$$p_m - c_m \geq \bar{\gamma}_s(p_s - c_m)$$

Otherwise the doctor would have incentive to recommend a serious treatment when the problem is actually minor. Doctor's profit is highly limited by this condition.

Furthermore, according to our assumption (1), as long as doctor can extract enough surplus from patients, it is cost efficient to recommend a lab test when the diagnosis is serious instead of providing a serious treatment directly.

Lemma 2 *(Truth-telling) In the optimal equilibrium without kickback, doctor truthfully reports the type and severity of problem.*

Doctor has no incentive to cheat on the type of problem as long as prices for the two types of problems are equal. We assume that the cost of treatment as well as lab test are symmetric, there is no incentive for the doctor to charge different prices. Otherwise, some price is not optimal. Then there is also no point to cheat on the type of problem. Hence we abstract from the subscripts for different labs in the notation in following discussion.

There may be incentives for the doctor to cheat only when the problem is minor. According to Lemma 1, doctor will not recommend a serious treatment directly. Since there is no kickback, when the problem is minor, there is no incentive for the doctor to first send the patients to a lab and then provides a minor treatment. In details, a patient accept the lab test recommendation as long as his expected payment is smaller than the loss from being untreated, that is

$$\begin{aligned} & \frac{\alpha\rho_s l_s + [(1-\alpha)(1-\theta)\rho_s + (1-\alpha)\theta\rho_m]l_m}{\alpha\rho_s + [(1-\alpha)(1-\theta)\rho_s + (1-\alpha)\theta\rho_m} \\ \geq \quad & p_L + \frac{\alpha\rho_s p_s + [(1-\alpha)(1-\theta)\rho_s + (1-\alpha)\theta\rho_m]p_m}{\alpha\rho_s + [(1-\alpha)(1-\theta)\rho_s + (1-\alpha)\theta\rho_m}. \end{aligned}$$

When $\rho_m = 0$, given the same prices of treatments and lab test, more patient will accept the recommendation for a lab test, resulting higher profit for the doctor.

Above Lemmas allow us to focus on the group of equilibria which doctor does not cheat on the equilibrium path. We construct an equilibrium in which doctor honestly reports the diagnosis about severity and type of problem, recommends a minor treatment when $d = m$ and refers patients to a laboratory when $d = s$. Patients accept doctor's recommendation if they expect the total fee to treat the problem is lower than their expected loss from being left untreated.

After a laboratory test is performed, the test result is verifiably revealed to both patient and doctor. Given this information, the problem should be treated properly. Acceptance rate for a serious treatment becomes one after a test. There will be trade if and only if $p_i \in [c_i, l_i]$.

To support this truth-telling equilibrium, we make assumption on the off-the-equilibrium path belief: if doctor recommends a serious treatment without first recommending a lab test, the patient believes that the problem is minor. This is the worst belief that the patients hold off the equilibrium path which will support the largest range of equilibria. Given this belief, patients with all types will reject the serious treatment.

Note that laboratories are just inactive players so there is no need to specify their strategies.

As we consider a no kickback environment, there is no need to specify aware and unaware patients. All patients form rational beliefs in the equilibrium. After diagnosis, patient will be informed about the problem *type* T and the diagnosis d with respect to the *severity*.

In the truth-telling equilibrium, when a patient faces a recommendation for a minor treatment ($t = m$), she believes that the problem is minor with probability one. As long as $p_m \leq l_m$, the patient will accept the treatment. When a serious treatment is recommended ($t = s$), she believes that the problem is minor with probability one. According to our assumption, $p_s \geq c_s > l_m$, the patient always rejects the recommendation for serious treatment regardless of her subjective loss from being untreated. When a lab test is recommended ($t = l$), she updates her belief about her problem according to Bayes' Rule:

$$\Pr(i = s | t = l, d = s) = \frac{\alpha}{\alpha + (1 - \alpha)(1 - \theta)} \equiv \tilde{\alpha}$$

A patient who is expected to bear a loss l_s from serious problem will accept if and only if the expected payment for treating the problem is smaller than the expected loss.

$$p_L + \tilde{\alpha} p_s + (1-\tilde{\alpha}) p_m \leq \tilde{\alpha} l_s + (1-\tilde{\alpha}) l_m$$

Given the same price, the more precise the diagnosis is, the larger the probability that a patient will accept the lab test as she holds a higher belief that she gets a serious problem. Define

$$l^n \equiv \min\{l_s \in [\underline{l}_s, \bar{l}_s] : p_L + \tilde{\alpha} p_s + (1-\tilde{\alpha}) p_m \leq \tilde{\alpha} l_s + (1-\tilde{\alpha}) l_m\},$$

as the smallest l_s among patients who accept lab test given the prices of the lab test and treatment. Then, given the strategy and price of treatments and test, patients adopt cutoff strategies to decide whether to accept the recommendation from doctor.

Define strategies of patients as $\gamma_t : [\underline{l}_s, \bar{l}_s] \times [c_i, l_i]_{i \in \{m,s\}} \times [c_L, \infty) \to \{0, 1\}$ for $\{m, s, L\}$. In details,

$$\gamma_s = 0$$

$$\gamma_m = \begin{cases} 0 & \text{if } p_m > l_m \\ 1 & \text{if } p_m \leq l_m \end{cases}$$

$$\gamma_L = \begin{cases} 0 & \text{if } l_s < l^n \\ 1 & \text{if } l_s \geq l^n \end{cases}$$

Since l_s follows uniform distribution[7], we can define average acceptance rate of lab test as following:

$$\bar{\gamma}_L = \int_{\underline{l}_s}^{\bar{l}_s} \gamma_L(l_s) dF(l_s) = \frac{\bar{l}_s - l^n}{\bar{l}_s - \underline{l}_s}$$

Given the beliefs and strategy of patients, we next want to show that there is no incentive for the doctor to deviate from the equilibrium strategy that we discribed above as long as all the prices are set properly.

Our discussion starts from the sub-recommendation game. After a patient visits, doctor gets diagnosis about the severity. Doctor can either recommend a minor treatment

[7] We assume l_s follows uniform distribution for simplicity. The result will not qualitatively change if l_s follows other more general distribution.

or send patient to a lab when $d = m$. If patient is sent to a lab, the outcome would be minor for sure (because $\Pr(i = m|d = m) = 1$). Then patient's problem will be treated at price p_m. Doctor's profit is the same if he recommends a minor treatment directly. Therefore, it is incentive compatible for the doctor to just honestly report the diagnosis and recommend a minor treatment when he already diagnosed the problem is minor. If $d = s$, it's dominated to charge p_m directly. Also, given patients' off-the-equilibrium-path belief, it is not optimal to charge p_s directly either. Therefore, doctor will refer the patient to a lab.

The only problem left for the doctor is to set the most profitable prices for treatments before patients come. We will see that the higher price a doctor charges for each treatment, the less patients accept recommendation for test. It is possible that all patients follows the recommendation as long as the prices of treatment is low enough. Compared with the standard credence goods model, in which positive rejection rate is necessary for a truth-telling equilibrium, since lab test provide information, it is not necessary for positive rejection rate in a truth-telling equilibrium.

The doctor's problem of maximizing the expected profit is given by

$$\max_{p_m,p_s,l^n} \pi^n = (1-\alpha)\,\theta(p_m - c_m) + [\alpha(p_s - c_s) + (1-\alpha)\,(1-\theta)(p_m - c_m)]\frac{\bar{l}_s - l^n}{\bar{l}_s - \underline{l}_s}$$

s.t

$$c_s \leq p_s \leq l^n \tag{2}$$

$$c_m \leq p_m \leq l_m \tag{3}$$

$$\underline{l}_s \leq l^n \leq \bar{l}_s \tag{4}$$

$$p_L + \tilde{\alpha}p_s + (1-\tilde{\alpha})p_m \leq \tilde{\alpha}l^n + (1-\tilde{\alpha})l_m \tag{5}$$

The first term of the profit funtion is from the patients who are diagnosed as minor problem and the second term is the profit earn from patients who are diagnosed as serious problem. A test is recommended and only a fraction of $\frac{\bar{l}_s - l^n}{\bar{l}_s - \underline{l}_s}$ patients accept the test.

Lemma 3 *(5) binds.*

If (5) does not bind, the doctor can increase profit by charging a higer price for either serious treatment or minor treatment without affecting the incentive of those patient who wants to accept the recommendation for lab test.

This lemma implies that there must be some patients in the market who are indifference between getting treated or left untreated. Since the measure of these patients are zero, it is without loss of generality to assume that they accept the lab test when they feel indifference. In the later dicussion with multiple type of patients, we'll show that the respective conditions are not always bind.

Then we are ready to charaterize the optimal equilibirum. The result is summerized in the following proposition.

Proposition 1 *When there is a competitive market for each type of laboratory and there is no money transfer from labs to the doctor, in the optimal equilibrium, the doctor always truthfully reports the diagnosis. The equilibrium price of treatments are as following:*

$$p_{mA} = p_{mB} = l_m$$

$$p_{sA} = p_{sB} = \begin{cases} \frac{\tilde{\alpha}(\bar{l}_s + c_s) - (1-\tilde{\alpha})(l_m - c_m) - c_L}{2\tilde{\alpha}} & \text{if } \ l^n > \underline{l}_s \\ \underline{l}_s - \frac{c_L}{\tilde{\alpha}} & \text{if } \ l^n = \underline{l}_s \end{cases}$$

Where

$$l^n = \max\{\underline{l}_s, \frac{\tilde{\alpha}(\bar{l}_s + c_s) - (1-\tilde{\alpha})(l_m - c_m) + c_L}{2\tilde{\alpha}}\}$$

According to Lemma 3, in the optimal equilibrium, prices satisfy

$$p_L + \tilde{\alpha} p_s + (1-\tilde{\alpha}) p_m = \tilde{\alpha} l^n + (1-\tilde{\alpha}) l_m \tag{6}$$

Due to the correct belief of patients in the equilibrium, all patients are expected to pay $\tilde{\alpha} l^n + (1-\tilde{\alpha}) l_m$ following a lab test, which only depends on l^n. As a result, as long as the adjustment of prices satisfying (6), it will not change the incentive for the patients to accept the lab test as well as the expected profit earned from each patient who is diagnosed as serious problem. A higher price of minor treatment increases the profit from patients who are diagonised as minor problem. As a result, treatments of minor problems are charged at monopoly price while treatments of serious problems are not.

To choose a proper price for serious problem, doctor faces a trade-off: a high price for serious problem results in high profit for individual patients while a low acceptance probability for the lab test. An optimal price of serious treatment should be smaller than the monopoly level.

Proposition **??** shows that under-provision of lab test occurs when $\frac{\tilde{\alpha}(\bar{l}_s+c_s)-(1-\tilde{\alpha})(l_m-c_m)+c_L}{2\tilde{\alpha}} > \underline{l}_s$. We can reorganize this condition as below:

$$\tilde{\alpha}(\underline{l}_s - c_s) + (1-\tilde{\alpha})(l_m - c_m) - c_L < \tilde{\alpha}(\bar{l}_s - \underline{l}_s)$$

The left-hand side is the expected profit of doctor when all patients are treated. The right-hand side is can be rewirte as $[\tilde{\alpha}\bar{l}_s + (1-\tilde{\alpha})l_m] - [\tilde{\alpha}\underline{l}_s + (1-\tilde{\alpha})l_m]$, the difference of patients' willing-to-pay for treatment between a highest-loss patient and the lowest patient. When patients have a very diversified subject feeling about the loss from a serious problem (large variation of distribution F), doctor can earn more from those high loss patients. In this case, doctor tends to charge a higher price and treat less patients.

Diagnosis precision also affect the acceptence rate of lab test in the optimal equilibrium. When both α and θ is small, the probability of wrong diagnosis $(1-\alpha)(1-\theta)$ becomes large. Doctor highly relies on test to get accurate diagnosis. Then lab tests are fully provided when necessary. Under-provision happens when either α or θ is large. Doctor's diagnosis becomes more precise. The result of a lab test is more likely to be a serious one. Some patients must bear a huge loss if they are left untreated, they are willing to pay much more than low-loss patients. As a result, setting a high price for serious treatment and only serving high valuation patients is more profitable for the doctor.

5 Steering advice through kickback

In practice, labs have incentive to pay doctor kickback in order to steer the recommendation of doctor. We consider contingent commissions from labs to doctor. If a lab and a doctor agree on the kickback, then whenever a patient is referred to a lab, kickback should be paid. This scenario resembles an "all-pay" auction, in that even if a lab offers less favorable kickback to doctor so that the recommendation is actually not steered, the lab must still pay kickback when the doctor refers any patient to it.

For the similar reason in previous section, it is not optimal for the doctor to recommend a serious treatment directly, not only because of the unverifiability of the treatment but also the cost-efficiency of using lab test. Furthermore, when kickback presents, doctor has larger incentive to recommend lab test in order to get extra payment from labs. We present a more general result in the following proposition. It holds for the environment with kickback and for any patients base. To keep our results clean and keep the essential intuition behind, from now on we focus on the interior solutions when we consider a herterogeneous base of patients.

Proposition 2 *Any equilibrium in which doctor recommends a serious treatment directly is not optimal for the doctor.*

5.1 Exploiting the unawareness of patients

To study the impact of kickback, we first consider naive patients who are unaware of the possibility of kickback from labs to doctor. From this section, we will add subscripts A and B, which refer to different labs (problems) since some asymmetry may happen.

Navie patients are unaware of the possibility that doctor may receive kickback from labs and ignore the incentives arising from commissions. Doctor tries to make use of the bounded rationality of patients. As a result, unnecessary or wrong test may appear in the equilibrium.

When the diagnosis of the problem is minor, with positive kickback, doctor now has incentive to recommend a lab test and gets kicback before treating the patients.

Due to the unverifiability of the treatment, doctor may also sometimes recommend a wrong type of lab test although he has perfect information about the type. This senerio may happen when the two labs offer different amount of kickbacks. For example, consider the situation that Lab A offers higher kickback than Lab B. When doctor diagnoses that the problem is minor and the type is B, doctor can cheat on both severity and type of problem and recommend a lab test A. Note that in this case doctor does not need information from lab test in order to figure out the most cost-efficient treatment. Then the doctor tends to refer the patient to a lab which gives him more kickbacks and finally provides a right type of minor treatment.

It's obvious that there is no point in cheating solely on the type of problem. Here we also exclude the uninteresting case that the doctor recommends a wrong test when he is not sure about the severity. As no useful information can be obtained through wrong test, the doctor must charge a minor treatment price but actually conduct a serious treatment according to our liability assumption.

As there are several ways for the doctor to exploit the naivity of patients, there will be several candidate equilibria for the optimal equilibrium. The following Lemma restricts our attention to the group of equilibria in which doctor cheats on both severity and type.

Lemma 4 *It is not optimal for the doctor to cheat solely on the severity.*

An intuitive explanation for above Lemma is that cheating on severity and type at the same time results in highest bias of unaware patients' belief on the probability of a serious problem. A higher bias makes patients willing to pay much more than they actually want to pay if they know the true likelihood. As we consider a monopoly doctor, he can extract more surplus from unaware patients in this way.

Now we construct an equilibrium that doctor and patients' equilibirum strategies and beliefs as following:

Doctor recommends a lab test A when $(d, T) \in \{(m, A), (m, B), (s, A)\}$, and recommends a lab test B when $(d, T) = (s, B)$. Doctor receives kickback k_A for each referral to lab A and k_B for each referral to lab B.

Naive patients accept direct recommendation of minor treatment as long as $p_{mT} \leq l_{mT}$, for $T \in \{A, B\}$. They reject direct recommendation for serious treatment. They accept the lab test as long as their expected payment for treating the problem is lower than the expected loss. Beliefs are specifed as below:

$$\begin{aligned}
\Pr(i &= s|t=l, d=s) = \tilde{\alpha} \\
\Pr(i &= m|t=m, d=m) = 1 \\
\Pr(i &= m|t=s, d=s) = 1
\end{aligned}$$

Again, to support this equilibrium, we make assumption off the equilibrium path as follows: when doctor recommends a serious treatment directly, patient believes that the problem is actually minor.

Prices and kickbacks are chosen to maximize the following profit function:

$$\begin{aligned}
\pi_{NK}^{u} &= \frac{1}{2}[\alpha(p_{sA}-c_s)+(1-\alpha)(p_{mA}-c_m)+k_A]\frac{\bar{l}_s - l_{NK_A}^{u}}{\bar{l}_s-\underline{l}_s} \\
&+\frac{1}{2}[(1-\alpha)\theta(p_{mA}-c_m+k_A)]\frac{\bar{l}_s - l_{NK_A}^{u}}{\bar{l}_s-\underline{l}_s} \\
&+\frac{1}{2}[\alpha(p_{sB}-c_s+k_B)+(1-\alpha)(1-\theta)(p_{mB}-c_m+k_B)]\frac{\bar{l}_s - l_{NK_B}^{u}}{\bar{l}_s-\underline{l}_s}
\end{aligned}$$

And $l_{KK\_T}^{u} \equiv \min\{l_s \in [\underline{l}_s, \bar{l}_s] : p_{LT} + \tilde{\alpha} p_{sT} + (1-\tilde{\alpha})p_{mT} \leq \tilde{\alpha} l_s + (1-\tilde{\alpha}) l_m\}$. Again, apart from the participation conditions parallelize to (2) to (4), when the problem is minor, kickback should be self-enforcing such that the doctor want to earn the kickback at the price that he sets for the treatment.

$$\frac{\bar{l}_s - l_{KK\_A}^{u}}{\bar{l}_s-\underline{l}_s}(k_A+p_{mA}-c_m) \geq \frac{\bar{l}_s - l_{KK\_B}^{u}}{\bar{l}_s-\underline{l}_s}(k_B+p_{mB}-c_m) \tag{7}$$

$$\frac{\bar{l}_s - l_{KK\_A}^{u}}{\bar{l}_s-\underline{l}_s}(k_A+p_{mA}-c_m) \geq p_{mT}-c_m \quad T\in\{A,B\} \tag{8}$$

(7) ensures that doctor wants to make a wrong recommendation when $(d,T)=(m,B)$. (8) is the incentive compatible condition for the doctor to cheat on severity when diagnosis is minor.

The following proposition charaterize the optimal equilibrium for the doctor in the senario that kickback is allowed.

Proposition 3 *When patients are unware of kickback, among the equilibria that the doctor receives kickback from labs, there exist an equilibrium that doctor obtains highest profit,* *in which doctor only charges kickback from one lab. More specifically, the doctor will set kickback and treatment prices as following:*

$$p_{mA} = p_{mB} = l_m$$

$$\begin{aligned} p_{sA} &= c_s \\ p_{sB} &= l^u_{NK\_B} - \frac{c_L}{\tilde{\alpha}} \end{aligned}$$

$$\begin{aligned} k_A &= \tilde{\alpha}(l^u_{NK\_A} - c_s) - c_L \\ k_B &= 0 \end{aligned}$$

Where

$$l^u_{NK\_A} = \max\{\underline{l}_s, \frac{\tilde{\alpha}(\bar{l}_s + c_s) - \frac{(1-\alpha)(1+\theta)}{1+(1-\alpha)\theta}(l_m - c_m) + c_L}{2\tilde{\alpha}}\}$$

$$l^u_{NK\_B} = \max\{\underline{l}_s, \frac{\tilde{\alpha}(\bar{l}_s + c_s) - (1-\tilde{\alpha})(l_m - c_m) + c_L}{2\tilde{\alpha}}\}$$

Although we assume that labs are symmetric in their test precision as well as cost of test, doctor still prefers an asymmetric kickback plan. When there is kickback, doctor always recommends the patients to the lab regardless of the diagnosis. Naive patients (weakly) over-estimate the probability of serious problem when they faces a recommendation for a lab test. In this case, doctor has incentive to extract all surplus from serious problem through kickback. In the equilibrium, offering serious treatment of type A problem itself results in zero profit. Compared with receiving kickbacks from both labs, one way for the doctor to improve profit is to keep the total amount of patients who accept lab test fixed, while increase the fraction of minor problems. By recommend all the minor problem to one lab, and lower the kickback from that lab, more minor problem patients will conduct lab test.

For the lab test without kickback (lab test B), the acceptence rate of lab test is same as in the no kickback case. This is because patients hold same beliefs about the probability of serious problem when they are recommended a lab test and the belief is correct. There are multiple equilibria with positive kickback resulting in the same profit after test B. We just choose the one with zero kickback.

Corollary 1 *When patients are unaware of kickback, kickback increases the acceptance rate of lab test from* Lab A *and results in more serious problem being treated.*

This result is straightforward following the propostion 3. When there is kickback, as doctor can earn more by exploiting naivety of the patients, and receives kickback from labs, he wants to include more patients into lab test.

We should also notice that although the acceptance rate of lab test is increased, the patients with minor problems being treated decreases. Patients with a low l_s and diagnosed as minor problems are left untreated.

5.2 Patients' Awareness

Suppose the market consists of a mixed population of unaware patients and aware patients. The fraction of patients being aware of kickback is ω. Aware patients understand that doctor may charge kickbacks from labs. Denote that aware patients' belief about kickback from the two labs are $\{\hat{k}_A, \hat{k}_B\}$. In the equilibrium, aware paitents hold correct belief about kickback, $\hat{k}_A = k_A$, $\hat{k}_B = k_B$. When a lab test is recommended, we denote the updated belief about probability of serious problem as $\{\hat{\alpha}^a_A, \hat{\alpha}^a_B\}$. Aware patient accepts a recommendation for a lab test $T \in \{A, B\}$ if

$$p_{LT} + \hat{\alpha}^a_T p_{sT} + (1 - \hat{\alpha}^a_T) p_{mT} \leq \hat{\alpha}^a_T l_s + (1 - \hat{\alpha}^a_T) l_m$$

Then she can evaluate the expected acceptence rate of lab test among all patients,

$$\bar{\gamma}_{LT} = \frac{\bar{l}_s - \hat{l}_{sT}}{\bar{l}_s - \underline{l}_s} \quad T \in \{A, B\}$$

$$\hat{l}_{sT} = \omega \hat{l}^a_{sT} + (1 - \omega) \hat{l}^u_{sT}$$

Aware patients' beliefs $\{\hat{k}_A, \hat{k}_B, \hat{\alpha}^a_A, \hat{\alpha}^a_B\}$ should be consistent with the equilibrium strategy of the doctor both on and off the equilibrium path and follows Bayes rule.

i) If $\{\hat{k}_A, \hat{k}_B, \hat{\alpha}^a_A, \hat{\alpha}^a_B\}$ satisfy

$$\hat{\alpha}^a_A = \frac{\alpha}{1 + (1 - \alpha)\theta}, \hat{\alpha}^a_B = \frac{\alpha}{1 - (1 - \alpha)\theta}$$

$$\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(\hat{k}_A + p_{mA} - c_m) > \frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s}(\hat{k}_B + p_{mB} - c_m)$$
$$\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(\hat{k}_A + p_{mA} - c_m) \geq p_{mA} - c_m$$

aware patients believe that doctor cheats on both severity and type of the problem and refers the patient to lab A when the diagnosis is minor problem B. Then

ii) If $\{\hat{k}_A, \hat{k}_B, \hat{\alpha}_A^a, \hat{\alpha}_B^a\}$ satisfy:

$$\hat{\alpha}_A^a = \hat{\alpha}_B^a = \alpha$$
$$\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(\hat{k}_A + p_{mA} - c_m) = \frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s}(\hat{k}_B + p_{mB} - c_m) \geq p_{mT} - c_m\ ,\ T \in \{A, B\}$$

aware patients believe that doctor only cheats on severity, and refers the patient to the lab when diagnosis is minor.

iii) If Then $\{\hat{k}_A, \hat{k}_B, \hat{\alpha}_A^a, \hat{\alpha}_B^a\}$ satisfy:

$$\hat{\alpha}_A^a = \hat{\alpha}_B^a = \frac{\alpha}{1 - (1-\alpha)\theta}$$
$$\hat{k}_A = \hat{k}_B = 0$$

$$\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(\hat{k}_A + p_{mA} - c_m) \leq p_{mT} - c_m,\ T \in \{A, B\}$$

aware patients believe that doctor tells the truth.

5.3 Influence of aware patients

In this section, we focus on an equilibrium in which doctor cheats on both severity and type in the equilibrium as before. More specifically, the doctor charges different kickbacks from Lab A and Lab B. He recommends all patients who are diagnosed as minor problem to Lab A and recommend patients who are diagnosed as serious problem to conduct lab test according to their true type. As we have shown, this is also the most profitable way to exploit naivity of patients. To support this equilibrium, again we assume the worst belief following a recommendation for serious treatment: patients believe that the problem is minor with probability one.

Given the beliefs specialized above, doctor has no incentive to deviate from the equilibrium in the recommendation subgame.

Then we only need to solve the pricing problem at the most beginning. Doctor's expected profit is given by:

$$\begin{aligned}
\pi^a &= \frac{1}{2}[\alpha(p_{sA}-c_s)+(1-\alpha)(1+\theta)(p_{mA}-c_m)+(1+(1-\alpha)\theta)k_A]\frac{\bar{l}_s-\hat{l}_{sA}}{\bar{l}_s-\underline{l}_s} \\
&+\frac{1}{2}[\alpha(p_{sB}-c_s)+(1-\alpha)(1-\theta)(p_{mB}-c_m)+(1-(1-\alpha)\theta)k_B]\frac{\bar{l}_s-\hat{l}_{sB}}{\bar{l}_s-\underline{l}_s}
\end{aligned}$$

subject to

$$\begin{aligned}
\hat{l}_{sA} &\geq p_{sT} \geq c_s \\
l_m &\geq p_{mT} \geq c_m \\
\bar{l}_s &\geq \hat{l}_{sA} \geq \underline{l}_s
\end{aligned}$$

$$\begin{aligned}
\frac{\bar{l}_s-\hat{l}_{sA}}{\bar{l}_s-\underline{l}_s}(k_A+p_{mA}-c_m) &> \frac{\bar{l}_s-\hat{l}_{sB}}{\bar{l}_s-\underline{l}_s}(k_B+p_{mB}-c_m) \\
\frac{\bar{l}_s-\hat{l}_{sA}}{\bar{l}_s-\underline{l}_s}(k_A+p_{mA}-c_m) &\geq p_{mA}-c_m
\end{aligned}$$

$$p_{LT}+\hat{\alpha}_T^a p_{sT}+(1-\hat{\alpha}_T^a)p_{mT} \leq \hat{\alpha}_T^a l_s+(1-\hat{\alpha}_T^a)l_m \tag{9}$$

$$p_{LT}+\hat{\alpha}_T^u p_{sT}+(1-\hat{\alpha}_T^u)p_{mT} \leq \hat{\alpha}_T^u l_s+(1-\hat{\alpha}_T^u)l_m \tag{10}$$

$$p_{LT}=k_T+c_L, T\in\{A,B\}$$

$$\hat{k}_T=k_T,\ T\in\{A,B\}$$

(9) and (10) may not bind at the same time in the optimal equilibrium. Then it can be the case that all aware patients get positive surplus. The result of the optimization problem is presented in the appendix. Here we show some comparative static results[8].

[8] We focus on the interior solution of the optimization problem above, some corner solution will be discussed in the appendix.

Proposition 4 *In the optimal equilibrium with kickback from labs to doctor, as more patients being aware of kickback,*

1) fewer kickback is received by the doctor;

2) the acceptence rate of lab test A is lower when cost of lab test is small, while acceptence rate of lab test B is not affected;

3) treatment prices are not affected.

When there are more patients being aware of kickback, the average belief on the problem's severity is less biased. There is less room for the doctor to make more profit through kickback. Patients expect to pay less to treat the problem following a lab test. As a results, given the same price of treatment, fewer patients will accept the recommendation for a lab test. Doctor tries to lower kickback a little bit to comprimise the effect of losing paiteints cause by more patients being aware.

6 Policy implications and Social Welfare

6.1 *Educating Patients*

There are a lot of methods which can improve patients' awareness about hidden payment between doctor and labs. For example, public lectures and advertisements. These methods can raise social awareness of kickback by making more people aware of this problem.

When cost of lab test is small, the existence of aware patients will lower the kickback. $\hat{l}_{sA} < \hat{l}^{u}_{NK\_A}$ and $\hat{l}_{sB} = \hat{l}^{u}_{NK\_B}$ implies that more patients with serious problems are being treated in the equilibrium with the existence of aware patients. As more patients becomes aware of kickback, the equilibrium level of kickback decreases. Aware patients expect a lower probability that the problem is serious following the recommendation of a lab test. As a result, given the same prices, fewer aware patients are willing to accept a lab test than naive patients. To mitigate the situation that rejection of lab test increases as more patients becomes aware, doctor will lower the kickback a little bit such that not too many aware patient go away. Awareness of some patients can protect the other naive patient from being exploit to some extent.

Kickback exists only when there is not too many aware patients in the market. As we'll show in Proposition 5, in the extreme case when $\omega = 1$, it is not profitable for the doctor to charge any kickback.

Lemma 5 *When all patients hold correct belief in the equilibrium, that is $\omega = 1$, doctor can not earn more than in the case of no kickback.*

When patients have correct belief in the equilibrium, there is no way for the doctor to earn more profit though kickback.

Proposition 5 *As more patients in the market becomes aware, it is less profitable for the doctor to charge kickback. There exsit some $\hat{\omega} \in [0, 1]$, such that when $\omega \geq \hat{\omega}$, doctor won't charge kickback.*

When $\omega < \hat{\omega}$, doctor still charge kickback from labs. Social welfare in the optimal equilibrium is

$$\begin{aligned} W^k &= \frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(-(1+(1-\alpha)\theta)c_L - \alpha c_s - (1-\alpha)(1+\theta)c_m) \\ &\quad + \frac{1}{2}\int_{\underline{l}_s}^{\hat{l}_{sA}} (-\alpha l_s - (1-\alpha)(1+\theta)l_m)dF(l_s) \\ &\quad \frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s}(-(1+(1-\alpha)\theta)c_L - \alpha c_s - (1-\alpha)(1-\theta)c_m) \\ &\quad + \frac{1}{2}\int_{\underline{l}_s}^{\hat{l}_{sB}} (-\alpha l_s - (1-\alpha)(1-\theta)l_m)dF(l_s) \end{aligned}$$

It's easy to check that $\frac{\partial W^k}{\partial \hat{l}_{sA}} < 0$ and $\frac{\partial W^k}{\partial \hat{l}_{sB}} < 0$[9] , a higher acceptence rate of the lab test improves social welfare. $\hat{l}_{sA}$ is decreasing in ω. Then when more patients becomes aware of kickback, social welfare as well as patient surplus can be improved. But when $\omega > \hat{\omega}$, the result can be opposite. When the fraction of aware patients is large enough,

[9]

$$\begin{aligned} \frac{\partial W^k}{\partial l_{sA}} &= \frac{1}{2}\frac{-1}{\bar{l}_s - \underline{l}_s}(-(1+(1-\alpha)\theta)c_L - \alpha c_s - (1-\alpha)(1+\theta)c_m) + \frac{1}{2}\frac{1}{\bar{l}_s - \underline{l}_s}(-\alpha l_{sA} - (1-\alpha)(1+\theta)l_m) \\ &= \frac{1}{2}\frac{1}{\bar{l}_s - \underline{l}_s}[-\alpha l_{sA} - (1-\alpha)(1+\theta)l_m + (1+(1-\alpha)\theta)c_L + \alpha c_s + (1-\alpha)(1+\theta)c_m] < 0 \end{aligned}$$

it is not profitable for the doctor to charge kickback anymore. In this case, educating patients have same effect as prohibiting the kickback directly. We will discuss the effect of banning kickback in the following analysis and show how it can hurt social welfare.

6.2 *Banning Kickback*

According to our assumption on the cost of treatments and lab test, using a lab test when the diagnosis is imperfect is costly efficient. The first best solution is to treat directly when diagnosis is minor and conduct a lab test before a treatment when the diagnosis is serious. Hence, when there is no kickback, the more patients being treated, the higher the social welfare. When there is no kickback, social welfare is caculated as below:

$$\begin{aligned} W^n &= -(1-\alpha)\theta c_m + \frac{\bar{l}_s - l^n}{\bar{l}_s - \underline{l}_s}(-(\alpha + (1-\alpha)(1-\theta))c_L - \alpha c_s - (1-\alpha)(1-\theta)c_m) \\ &\quad + \int_{\underline{l}_s}^{l^n} [-\alpha l_s - (1-\alpha)(1-\theta)l_m] dF(l_s) \end{aligned}$$

The first two terms are the cost of treating patients. The third term is the loss of patient being untreated.

It is obvious that lower cost for treatments and test will improve social welfare not only by lower the cost of treatment, but also raise the acceptence rate in the optimal equilibrium[10]. A more diversified patient loss will hurt social welfare since high-loss patients attract doctor to set a higher price. As a result, very little patients get treated.

When there is no under-provision problem in no kickback case, there is no need to ban kickback since doctor can not earn more through kickback. The reason is simple, when

---

$$\begin{aligned} \frac{\partial W^k}{\partial l_{sB}} &= \frac{1}{2}\frac{-1}{\bar{l}_s - \underline{l}_s}(-(1+(1-\alpha)\theta)c_L - \alpha c_s - (1-\alpha)(1-\theta)c_m) + \frac{1}{2}\frac{1}{\bar{l}_s - \underline{l}_s}(-\alpha l_{sB} - (1-\alpha)(1-\theta)l_m) \\ &= \frac{1}{2}\frac{1}{\bar{l}_s - \underline{l}_s}[-\alpha l_{sB} - (1-\alpha)(1-\theta)l_m + (1+(1-\alpha)\theta)c_L + \alpha c_s + (1-\alpha)(1-\theta)c_m] < 0 \end{aligned}$$

[10]

$$\begin{aligned} \frac{\partial W^n}{\partial l^n} &= \frac{-1}{\bar{l}_s - \underline{l}_s}(-(\alpha + (1-\alpha)(1-\theta))c_L - \alpha c_s - (1-\alpha)(1-\theta)c_m) + \frac{1}{\bar{l}_s - \underline{l}_s}(-\alpha l^n - (1-\alpha)(1-\theta)l_m) \\ &= \frac{1}{\bar{l}_s - \underline{l}_s}[-\alpha l^n - (1-\alpha)(1-\theta)l_m + (\alpha + (1-\alpha)(1-\theta))c_L + \alpha c_s + (1-\alpha)(1-\theta)c_m] < 0 \end{aligned}$$

all patients get treated, the expected payment after a lab test are the same no matter whether there is kickback. Doctor need to bare the wasted cost of unnecessary and wrong lab test. As a result, doctor earns less if he recommamds any unnecessary or wrong test.

However, when lab test is **under-provided** by no kickback labs, it is not clear how kickback will affect the social welfare. As we have shown, kickback can improve acceptence rate of lab test, resulting in more serious problem patients being treated in the equilibrium. At the same time, the minor problem being treated may decreases when doctor cheats on the severity of problem. Some costs are also wasted on the wrong tests and unnecessary test. More patients being treated means more cost is wasted. These effects have the opposite impact on the social welfare.

Compare social welfare in the case with no kickback and kickback,When the cost of lab test is small, and kickback sufficiently alleviate the problem of under-provision of lab test and treatment, banning kickback aggravates the problem of under-provison and hurt social welfare[11].

7 Conclusion

Due to the unverifiability of treatments, patients can not fully trust doctor in his recommendation. As Labarotories provide hard evidence for the diagnosis, as long as it costs not much, it would be benefitial for the socioty as well as doctor to conduct test when the diagnosis is not perfect. Kickback from labs can easily ruin these benefits and steer the advice of the doctor. Patients are directed to unnecessary as well as more expensive tests. Patients pay much more then they expect if there are unaware of kickback. In a market with mixed population of aware and unaware patients, some patients being aware of kickback can protect all patients from being over-charged too much. As increasing patients being aware of kickback, fewer kickback is taken from lab by the doctor. Banning kickack can have unintended consequences.

[11] Here, in the case with kickback, we assume that the fraction of aware patients is small enough such that doctor still wants to charge kickback from labs.

References

[1] Darby, Michael R., and Edi Karni. "Free competition and the optimal amount of fraud." *The Journal of law and economics* 16.1 (1973): 67-88.

[2] Wolinsky, Asher. "Competition in a market for informed experts' services." *The RAND Journal of Economics* (1993): 380-398.

[3] Fong, Yuk-fai. "When do experts cheat and whom do they target?." *RAND Journal of Economics* (2005): 113-130.

$$
\begin{aligned}
W^n - W^k &= -(1-\alpha)\theta c_m - \frac{1}{2}\frac{l^n - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[-(\alpha + (1-\alpha)(1-\theta))c_L - \alpha c_s - (1-\alpha)(1-\theta)c_m] + \frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(1-\alpha)\theta c_L \\
&\quad + \frac{1}{2}\int_{\hat{l}_{sA}}^{l^n}[-\alpha l_s - (1-\alpha)(1-\theta)l_m]dF(l_s) - \int_{\underline{l}_s}^{\hat{l}_{sA}} - (1-\alpha)\theta)l_m dF(l_s) \\
&= +\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(1-\alpha)\theta c_L - (1-\alpha)\theta c_m + (1-\alpha)\theta[\frac{\bar{l}_s - l_{sA}}{\bar{l}_s - \underline{l}_s}c_m + \frac{\hat{l}_{sA} - \underline{l}_s}{\bar{l}_s - \underline{l}_s}l_m] \\
&\quad - \frac{1}{2}\frac{l^n - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[-(\alpha + (1-\alpha)(1-\theta))c_L - \alpha c_s - (1-\alpha)(1-\theta)c_m] \\
&\quad + \frac{1}{2}\frac{1}{\bar{l}_s - \underline{l}_s}[-\frac{\alpha}{2}(l_s)^2 - (1-\alpha)(1-\theta)l_m l_s]|_{l_{sA}}^{l^n} \\
&= \frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(1-\alpha)\theta c_L - (1-\alpha)\theta c_m - \frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(1-\alpha)\theta c_m \\
&\quad - \frac{1}{2}\frac{l^n - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[-(\alpha + (1-\alpha)(1-\theta))c_L - \alpha c_s - (1-\alpha)(1-\theta)c_m] \\
&\quad + \frac{1}{2}\frac{1}{\bar{l}_s - \underline{l}_s}[-\frac{\alpha}{2}(l^n)^2 - (1-\alpha)(1-\theta)l_m l^n + \frac{\alpha}{2}(\hat{l}_{sA})^2 + (1-\alpha)(1-\theta)l_m\hat{l}_{sA}] \\
&= \frac{\hat{l}_{sA} - \underline{l}_s}{\bar{l}_s - \underline{l}_s}(1-\alpha)\theta(l_m - c_m) + \frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(1-\alpha)\theta c_L \\
&\quad - \frac{1}{2}\frac{l^n - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[\frac{\alpha}{2}(l^n + \hat{l}_{sA}) + (1-\alpha)(1-\theta)l_m - (\alpha + (1-\alpha)(1-\theta))c_L - \alpha c_s - (1-\alpha)(1-\theta)c_m] \\
&= \frac{\hat{l}_{sA} - \underline{l}_s}{\bar{l}_s - \underline{l}_s}(l_m - c_m) + \frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(1-\alpha)\theta c_L \\
&\quad - \frac{1}{2}\frac{l^n - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[\alpha(\frac{l^n + \hat{l}_{sA}}{2} - c_s) + (1-\alpha)(1-\theta)(l_m - c_m) - (\alpha + (1-\alpha)(1-\theta))c_L]
\end{aligned}
$$

The first term of the last equation is the social welfare loss from kickback: some patients who are diagnosed as minor problem are not treated. The second term is the loss from wasted cost on the lab test. The third term is the social welfare gain from kickback. More patients who are diagnosed as serious problem are being treated. As long the the first term and the second term is small enough, banning kickback will hurt social welfare.

[4] Emons, Winand. "Credence goods monopolists." *International Journal of Industrial Organization* 19.3-4 (2001): 375-389.

[5] Emons, Winand. "Credence goods and fraudulent experts." *The RAND Journal of Economics* (1997): 107-119.

[6] Dulleck, Uwe, and Rudolf Kerschbamer. "On doctors, mechanics, and computer specialists: The economics of credence goods." *Journal of Economic literature* 44.1 (2006): 5-42.

[7] Liu, Ting. "Credence goods markets with conscientious and selfish experts." *International Economic Review* 52.1 (2011): 227-244.

[8] Hilger, Nathaniel G. "Why Don't People Trust Experts?." *The Journal of Law and Economics* 59.2 (2016): 293-311.

[9] Pauly, Mark V. "The ethics and economics of kickbacks and fee splitting." *The Bell Journal of Economics* (1979): 344-352.

[10] Inderst, Roman, and Marco Ottaviani. "Competition through commissions and kickbacks." *American Economic Review* 102.2 (2012): 780-809.

[11] Inderst, Roman. "Regulating commissions in markets with advice." *International Journal of Industrial Organization* 43 (2015): 137-141.

[12] Inderst, Roman, and Marco Ottaviani. "How (not) to pay for advice: A framework for consumer financial protection." Journal of Financial Economics 105.2 (2012): 393-411.

[13] Pitchik, Carolyn, and Andrew Schotter. "Honesty in a model of strategic information transmission." *The American Economic Review* 77.5 (1987): 1032-1036.

8 Appendix A: Proofs of lemmas and propositions

Proof of Lemma 1. Without loss of generality, suppose the optimal equilibrium strategies of the doctor specify a price list $\{p_s^*, p_m^*\}$ and, for the recommendation subgame posting of every price list $\{p_s^*, p_m^*\}$, the probabilities $\{\beta_d^*, \rho_d^*, 1-\beta_d^*-\rho_d^*\}$ for each diagnosis $d \in \{m, s\}$. The equilibrium strategy of labs is a the price of lab test p_L^*. A mixed strategy of the patients specify a list of mapping $\{\gamma_t^*(l_m, l_s), 1-\gamma_t^*(l_m, l_s)\}$ for $t \in \{m, s, l\}$ and $p_t \in \{p_s^*, p_m^*, p_L^*\}$.

$$\bar{\gamma}_s^* = \int_{\underline{l}_s}^{\bar{l}_s} \gamma_t^*(l_m, l_s) dF(l_s)$$

$\beta_s^* > 0$ implies $\bar{\gamma}_s^* > 0$. When diagnosis is serious, it is strictly dominated to recommend a minor treatment, $\rho_s^* + \beta_s^* = 1$, then $\rho_s^* \in [0, 1)$. As there is no kickback, $\rho_m^* = 0$ for the reason that $t = l$ is a weakly dominated strategy by $t = m$ when $d = m$.

Patients' strategy:

1) $\gamma_m^* = 1$ if $p_m \leq l_m$;

2) $\gamma_s^*(l_m, l_s) = 1$ if

$$\frac{\alpha\beta_s^* l_s + [(1-\alpha)(1-\theta)\beta_s^* + (1-\alpha)\theta\beta_m^*] l_m}{[\alpha + (1-\alpha)(1-\theta)]\beta_s^* + (1-\alpha)\theta\beta_m^*} \geq p_s^*$$

Since $\beta_s^* > 0$

$$\frac{\alpha l_s + [(1-\alpha)(1-\theta) + (1-\alpha)\theta\frac{\beta_m^*}{\beta_s^*}] l_m}{\alpha + (1-\alpha)(1-\theta) + (1-\alpha)\theta\frac{\beta_m^*}{\beta_s^*}} \geq p_s^*$$

By assumption $\alpha\bar{l}_s + (1-\alpha) l_m < c_s$ that $\beta_m^* < \beta_s^* \leq 1$,then we have $0 \leq \frac{\beta_m^*}{\beta_s^*} < 1$.

3) $\gamma_L^*(l_m, l_s) = 1$ if

$$\frac{\alpha l_s + (1-\alpha)(1-\theta) l_m}{\alpha + (1-\alpha)(1-\theta)} \geq p_L + \frac{\alpha p_s^* + (1-\alpha)(1-\theta) p_m^*}{\alpha + (1-\alpha)(1-\theta)}$$

$$\begin{aligned} \widehat{l}_s^L &= \frac{[\alpha + (1-\alpha)(1-\theta)] p_L + \alpha p_s^* + (1-\alpha)(1-\theta)(p_m^* - l_m)}{\alpha} \\ &= \frac{[\alpha + (1-\alpha)(1-\theta)] c_L + \alpha p_s^*}{\alpha} \end{aligned}$$

$$\bar{\gamma}_L^* = \frac{\bar{l}_s - \widehat{l}_s^L}{\bar{l}_s - \underline{l}_s}$$

$$\bar{\gamma}_s^* = \frac{\bar{l}_s - \widehat{l}_s^s}{\bar{l}_s - \underline{l}_s}$$

Doctor's strategy:

$\beta_s^* > 0$ implies

$$\bar{\gamma}_s^*(p_s^* - c_s) \geq \bar{\gamma}_L^* \frac{\alpha(p_s^* - c_s) + (1-\alpha)(1-\theta)(p_m^* - c_m)}{\alpha + (1-\alpha)(1-\theta)} \tag{11}$$

$\beta_m^* < 1$ implies

$$p_m^* - c_m \geq \bar{\gamma}_s^*(p_s^* - c_m) \tag{12}$$

(12) and $p_s^* \geq c_s > l_m \geq p_m^*$ imply $\bar{\gamma}_s^* < 1$.

Then in this equilibrium, there exist some patient whose loss is $(l_m, \hat{l}_s)$, and is indifference between accept a serious treatment and reject.

$$\frac{\alpha\hat{l}_s + [(1-\alpha)(1-\theta) + (1-\alpha)\theta\frac{\beta_m^*}{\beta_s^*}]l_m}{\alpha + (1-\alpha)(1-\theta) + (1-\alpha)\theta\frac{\beta_m^*}{\beta_s^*}} = p_s^*$$

For the patients who have $l_s \geq \hat{l}_s$, will accpet the serious treatment, hence the acceptence rate of serious treatment is given by $\bar{\gamma}_s^* = \frac{\bar{l}_s - \hat{l}_s}{\bar{l}_s - \underline{l}_s}$. The equilibrium profit of doctor is

$$\begin{aligned}\pi^* &= (1-\alpha)\theta[(1-\beta_m^*)(p_m^* - c_m) + \beta_m^*\bar{\gamma}_s^*(p_s^* - c_m)] \\ &\quad +[\alpha + (1-\alpha)(1-\theta)]\bar{\gamma}_s^*(p_s^* - c_s)\end{aligned}$$

Now we can construct an equilibrium in which doctor always recommend a lab test ($\beta_s = 0$) when $d = s$, and recommend a minor treatment when $d = m$, and set price of serious treatment as $\tilde{p}_s$ so that accpetence rate of lab test is $\bar{\gamma}_L = \frac{\bar{l}_s - \hat{l}_s}{\bar{l}_s - \underline{l}_s} = \bar{\gamma}_s^*$

$$\frac{\alpha\hat{l}_s + (1-\alpha)(1-\theta)l_m}{\alpha + (1-\alpha)(1-\theta)} = p_L + \frac{\alpha\tilde{p}_s + (1-\alpha)(1-\theta)p_m^*}{\alpha + (1-\alpha)(1-\theta)}$$

To support this equilibrium, we only need to assume the off-equilibrium path belief as following: when patient observes a serious treatment recommendation, she believes that the probability that the problem is one. The profit that a doctor can get from this constructive equibilibrium is

$$\tilde{\pi} = (1-\alpha)\theta(p_m^* - c_m) + [\alpha + (1-\alpha)(1-\theta)]\bar{\gamma}_L\frac{\alpha(\tilde{p}_s - c_s) + (1-\alpha)(1-\theta)(p_m^* - c_m)}{\alpha + (1-\alpha)(1-\theta)}$$

$$\begin{aligned}\tilde{\pi} - \pi^* &\geq [\alpha + (1-\alpha)(1-\theta)]\frac{\bar{l}_s - \hat{l}_s}{\bar{l}_s - \underline{l}_s}[\frac{\alpha(\tilde{p}_s - c_s) + (1-\alpha)(1-\theta)(p_m^* - c_m)}{\alpha + (1-\alpha)(1-\theta)} - (p_s^* - c_s)] \\ &\geq [\alpha + (1-\alpha)(1-\theta)]\frac{\bar{l}_s - \hat{l}_s}{\bar{l}_s - \underline{l}_s}[\frac{\alpha(\hat{l}_s - c_s) + (1-\alpha)(1-\theta)(l_m - c_m)}{\alpha + (1-\alpha)(1-\theta)} \\ &\quad -p_L - (\frac{\alpha\hat{l}_s + (1-\alpha)(1-\theta)l_m}{\alpha + (1-\alpha)(1-\theta)} - c_s)] \\ &= [\alpha + (1-\alpha)(1-\theta)]\frac{\bar{l}_s - \hat{l}_s}{\bar{l}_s - \underline{l}_s}[\frac{(1-\alpha)(1-\theta)(c_s - c_m)}{\alpha + (1-\alpha)(1-\theta)} - p_L]\end{aligned}$$

Since labs are competitive, by assumption that $\frac{(1-\alpha)(1-\theta)}{\alpha+(1-\alpha)(1-\theta)}(c_s-c_m) > c_L$, we have $\tilde{\pi} > \pi^*$. Then we can conclude that any equilibrium with $\beta_s^* > 0$ is not optimal. ■

Proof of Proposition 1. (5) binds. Otherwise, the doctor can increase either p_s or p_m so that the expected profit increases.

$$(5) \text{ binds} \Longrightarrow p_s = l^n + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_m) - \frac{c_L}{\tilde{\alpha}}$$

By assumption $c_s < \underline{l}_s - \frac{c_L}{\tilde{\alpha}}$, $p_s > c_s$ in the optimal solution.

We can rewrite doctor's profit as

$$\begin{aligned}\pi^d &= (1-\alpha)\,\theta(p_m - c_m) \\ &\quad +[\alpha(l^n + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_m) - \frac{c_L}{\tilde{\alpha}} - c_s) + (1-\alpha)\,(1-\theta)(p_m - c_m)]\frac{\bar{l}_s - l^n}{\bar{l}_s - \underline{l}_s}\end{aligned}$$

The profit function is well defined: continuous and concave in l^n and p_m.

$$\begin{aligned}\frac{\partial \pi^d}{\partial p_m} &= (1-\alpha)\,\theta + \frac{\bar{l}_s - l^n}{\bar{l}_s - \underline{l}_s}[(1-\alpha)\,(1-\theta) - \alpha\frac{(1-\tilde{\alpha})}{\tilde{\alpha}}] \\ &= (1-\alpha)\,\theta > 0\end{aligned}$$

Then we have $p_m = l_m$.

$$\begin{aligned}\frac{\partial \pi^d}{\partial l^n} &= \alpha\frac{\bar{l}_s - l^n}{\bar{l}_s - \underline{l}_s} - [\alpha l^n - \alpha\frac{c_L}{\tilde{\alpha}} - \alpha c_s + (1-\alpha)\,(1-\theta)(l_m - c_m)]\frac{1}{\bar{l}_s - \underline{l}_s} \\ &= \frac{1}{\bar{l}_s - \underline{l}_s}[\alpha(\bar{l}_s - l^n) - \alpha l^n + \alpha\frac{c_L}{\tilde{\alpha}} + \alpha c_s - (1-\alpha)\,(1-\theta)(l_m - c_m)]\end{aligned}$$

If $\bar{l}_s - \frac{\tilde{\alpha}(\bar{l}_s - c_s) + (1-\tilde{\alpha})(l_m - c_m) - c_L}{2\tilde{\alpha}} > \underline{l}_s$,

$$\frac{\partial \pi^d}{\partial l^n} = 0 \Longrightarrow \alpha(\bar{l}_s - l^n) - \alpha l^n + \alpha\frac{c_L}{\tilde{\alpha}} + \alpha c_s - (1-\alpha)\,(1-\theta)(l_m - c_m) = 0$$

$$\begin{aligned}l^n &= \frac{\tilde{\alpha}\bar{l}_s + \tilde{\alpha}c_s + c_L - (1-\tilde{\alpha})(l_m - c_m)}{2\tilde{\alpha}} \\ &= \bar{l}_s - \frac{\tilde{\alpha}(\bar{l}_s - c_s) + (1-\tilde{\alpha})(l_m - c_m) - c_L}{2\tilde{\alpha}}\end{aligned}$$

$$\begin{aligned}p_s &= l^n - \frac{c_L}{\tilde{\alpha}} \\ &= \frac{\tilde{\alpha}\bar{l}_s + \tilde{\alpha}c_s - c_L - (1-\tilde{\alpha})(l_m - c_m)}{2\tilde{\alpha}}\end{aligned}$$

Otherwise the optimal solution is $l^n = \underline{l}_s$. ∎

Proof of Proposition 2. Consider a mixed strategy equilibrium. Suppose the optimal equilibrium strategy for the doctor is $\{\beta_{sT}^{T'}, \rho_{sT}^{T'}, \beta_{mT}^{T'}, \rho_{mT}^{T'}, p_{sT}, p_{mT}, k_T\}_{T,T' \in \{A,B\}}$ with $\beta_{sA}^{T'} > 0$ or $\beta_{mA}^{T'} > 0$. The equilibrium acceptence rate implied by the patients' equilibrium strategies are $\{\bar{\gamma}_{sT}, \bar{\gamma}_{mT}, \bar{\gamma}_{LT}\}$.

$\rho_{sA}^{B} = \rho_{sB}^{A} = 0$, since there is no need to recommend a wrong test when doctor actually need the information from the lab test. We exclude this type of equilibrium.

Step 1: Show $\beta_{mA}^{T'} < 1$.

Suppose $\beta_{mA}^{T'} = 1$ for some $T' \in \{A, B\}$, then given any recommendation for a direct serious treatment T', patients will reject due to our assumption: $\alpha \bar{l}_s + (1-\alpha)\, l_m < c_s$. As a result, $\bar{\gamma}_{sT'} = 0$. Then $\beta_{mA}^{T'} = 1$ is not an equilibrium strategy.

Step 2: Show $\beta_{sA}^{T'} < 1$.

Suppose $\beta_{sA}^{T'} = 1$ for some $T' \in \{A, B\}$. When $T' \neq A$, it means that doctor always recommends a wrong treatment when the problem serious. When $T' = A$, doctor does not lie on the type of problem.

$$p_{sT'} \leq \alpha_s \hat{l}_s + (1-\alpha_s) l_m$$

where

$$\begin{aligned}
\alpha_s &= \frac{\alpha(\beta_{sA}^{T'} + \beta_{sB}^{T'})}{[\alpha + (1-\alpha)(1-\theta)](\beta_{sA}^{T'} + \beta_{sB}^{T'}) + (1-\alpha)\,\theta(\beta_{mA}^{T'} + \beta_{mB}^{T'})} \\
&= \frac{\alpha}{[\alpha + (1-\alpha)(1-\theta)] + (1-\alpha)\,\theta \frac{\beta_{mA}^{T'} + \beta_{mB}^{T'}}{\beta_{sA}^{T'} + \beta_{sB}^{T'}}} \leq \frac{\alpha}{[\alpha + (1-\alpha)(1-\theta)]}
\end{aligned}$$

No lab test is adopted when the diagnosis is serious and problem is type A in the equilibrium. Now we can construct an equilibrium in which $k = 0$, doctor always recommend a lab test ($\beta_{sA}^{T'} = 0$) when $d = s, T = A$, and recommend a minor treatment when $d = m, T = A$, and set prices so that acceptence rate of lab test is $\hat{\gamma}_{LA} = \bar{\gamma}_{sT'} = \frac{\bar{l}_s - \hat{l}_s}{\bar{l}_s - \underline{l}_s}$. We keep the strategy regarding problem B the same as in the orginal eqilibrium.

$$\frac{\alpha \hat{l}_s + (1-\alpha)(1-\theta)\, l_m}{\alpha + (1-\alpha)(1-\theta)} = p_L + \frac{\alpha \tilde{p}_{sA} + (1-\alpha)(1-\theta)\, p_{mA}}{\alpha + (1-\alpha)(1-\theta)}$$

To support this equilibrium, we only need to specify the off-equilibrium-path belief as following: when patient observes a serious treatment recommendation, he believes that the problem is actually minor.

Since now we have $\beta_{mA}^{T'} < 1$, then [12]

$$p_{mA} - c_m \geq \bar{\gamma}_{sT'}(p_{sT'} - c_m)$$

The profit when problem is type A from this constructive equilibrium is

$$\tilde{\pi}^A = (1-\alpha)\,\theta(p_{mA} - c_m) + [\alpha + (1-\alpha)\,(1-\theta)]\frac{\bar{l}_s - \hat{l}_s}{\bar{l}_s - \underline{l}_s}\frac{\alpha(\tilde{p}_s - c_s) + (1-\alpha)\,(1-\theta)\,(p_{mA} - c_m)}{\alpha + (1-\alpha)\,(1-\theta)}$$

The profirt when problem is type A in the orginal eqilibrium

$$\pi^{A*} \leq (1-\alpha)\,\theta(p_{mA} - c_m) + [\alpha + (1-\alpha)\,(1-\theta)]\bar{\gamma}_{sT'}(p_{sT'} - c_s)$$

$$\begin{aligned}
\tilde{\pi}^A - \pi^{A*} &\geq [\alpha + (1-\alpha)\,(1-\theta)]\frac{\bar{l}_s - \hat{l}_s}{\bar{l}_s - \underline{l}_s}[\frac{\alpha(\tilde{p}_s - c_s) + (1-\alpha)\,(1-\theta)\,(p_m^* - c_m)}{\alpha + (1-\alpha)\,(1-\theta)} - (p_s^* - c_s)] \\
&\geq [\alpha + (1-\alpha)\,(1-\theta)]\frac{\bar{l}_s - \hat{l}_s}{\bar{l}_s - \underline{l}_s}[\frac{\alpha(\hat{l}_s - c_s) + (1-\alpha)\,(1-\theta)\,(l_m - c_m)}{\alpha + (1-\alpha)\,(1-\theta)} \\
&\quad - p_L - (\frac{\alpha\hat{l}_s + (1-\alpha)\,(1-\theta)\,l_m}{\alpha + (1-\alpha)\,(1-\theta)} - c_s)] \\
&= [\alpha + (1-\alpha)\,(1-\theta)]\frac{\bar{l}_s - \hat{l}_s}{\bar{l}_s - \underline{l}_s}[\frac{(1-\alpha)\,(1-\theta)\,(c_s - c_m)}{\alpha + (1-\alpha)\,(1-\theta)} - c_L]
\end{aligned}$$

Since labs are competitive, by assumption that $\frac{(1-\alpha)(1-\theta)}{\alpha+(1-\alpha)(1-\theta)}(c_s - c_m) > c_L$, we have $\tilde{\pi}^A > \pi^{A*}$. Contradiction.

Step 3: Show $\beta_{mA}^{T'} \in (0,1)$ or $\beta_{sA}^{T'} \in (0,1)$ is not optimal for the doctor. Consider an optimal equilibrium in which $\beta_{mA}^{T'} \in (0,1)$ or $\beta_{sA}^{T'} \in (0,1)$. As the treatments are unverifiable as long as the problem has been treated, and now direct recommendation of serious treamment with type T' is an on-the-equilibrium-path strategy, we must have

$$\bar{\gamma}_{sT'}(p_{sT'} - c_m) \leq p_{mT'} - c_m \leq l_m - c_m$$

And the acceptence rate of serious treatment is defined and restricted by following conditions:

$$\bar{\gamma}_{sT'} = \frac{\bar{l}_s - \hat{l}_s^s}{\bar{l}_s - \underline{l}_s}$$

[12] Given $\beta_{sA}^{T'} = 1$, when any lab test is recommended, patients should regard it as a unnecessary testing. In this case, they will reject.

$$p_{sT'} \leq \alpha_s \hat{l}_s + (1-\alpha_s) l_m$$

where

$$\begin{aligned}\alpha_s &= \frac{\alpha(\beta^{T'}_{sA}+\beta^{T'}_{sB})}{[\alpha+(1-\alpha)(1-\theta)](\beta^{T'}_{sA}+\beta^{T'}_{sB})+(1-\alpha)\,\theta(\beta^{T'}_{mA}+\beta^{T'}_{mB})} \\ &= \frac{\alpha}{[\alpha+(1-\alpha)(1-\theta)]+(1-\alpha)\,\theta\frac{\beta^{T'}_{mA}+\beta^{T'}_{mB}}{\beta^{T'}_{sA}+\beta^{T'}_{sB}}} \leq \frac{\alpha}{[\alpha+(1-\alpha)(1-\theta)]}\end{aligned}$$

When diagnosis is serious type A problem, $\beta^{T'}_{sA} \in (0,1)$ also implies that

$$\bar{\gamma}_{sT'}(p_{sT'}-c_s) = \bar{\gamma}_{LA}[k_A + \tilde{\alpha}(p_{sA}-c_s) + (1-\tilde{\alpha})(p_{mA}-c_m)]$$

The acceptence rate of a lab test A is given by:

$$\bar{\gamma}_{LA} = \frac{\bar{l}_s - \hat{l}^L_s}{\bar{l}_s - \underline{l}_s}$$

$$\hat{l}^L_s = \omega \hat{l}^{aL}_s + (1-\omega)\hat{l}^{uL}_s$$

For aware patients

$$p_{LA} + \alpha^a_L p_{sA} + (1-\alpha^{a*}_L) p_{mA} \leq \alpha^a_L \hat{l}^{aL}_s + (1-\alpha^a_L) l_m$$

$$\alpha^a_{LA} = \frac{\alpha\rho^A_{sA}}{[\alpha+(1-\alpha)(1-\theta)]\rho^A_{sA}+(1-\alpha)\,\theta(\rho^A_{mA}+\rho^B_{mA})}$$

For unaware patients

$$p_{LA} + \alpha^u_L p_{sA} + (1-\alpha^u_L) p_{mA} \leq \alpha^u_L \hat{l}^{uL}_s + (1-\alpha^u_L) l_m$$

$$\alpha^u_L = \frac{\alpha}{[\alpha+(1-\alpha)(1-\theta)]}$$

Next we want to construct another equilibrium in which when doctor's diagnosis is minor, he recommends a minor treatment directly. When diagnosis is serious, doctor always recommend a lab test. We keep the price $\hat{p}_{mA} = p_{mA}$. Then in the new equilibrium, we have $\hat{\beta}^{T'}_{sA} = \hat{\beta}^{T'}_{mA} = 0$, and $\hat{\rho}^A_{sA} = 1$, $\hat{\rho}^A_{mA} = 0$ (Since in this new equilibrium, there is no cheating, so we do not need to specify whether the patients are aware of kickback). We will show that this is an equilibrium and it gives doctor higher profit. If $\beta^{T'}_{mA} \in (0,1)$, then in the new equilibrium, when diagonsis is minor of type A, we offer an minor treatment directly at same price will not hurt profit and patients always accept the recommendation. Since

$$\bar{\gamma}_{sT'}(p_{sT'} - c_m) \leq p_{mA} - c_m \leq l_m - c_m$$

If $\beta_{sA}^{T'} \in (0, 1)$, when diagnosis is serious of type A, given that $\hat{\rho}_{sA}^{A} = 1$, $\hat{\rho}_{mA}^{A} = 0$ we have

$$\hat{\alpha}_{LA}^{a} > \alpha_{LA}^{a}$$

To keep the same price of lab test and the same acceptence rate of lab test, in the equilibrium, $\hat{p}_{sA} > p_{sA}$. As a result, the equilibrium profit is higher in this new equilibrium.

Next, we want to show that this is an equilibrium. To support this equilibrium, we specify the off-equilinrium path belief that the consumer believes that the when the doctor recommends a serious treatment directly, the problem is actually minor. Hence they reject the recommendation. Then we will see that when diagnosis is minor, there is no incentive for the doctor to deviate as doctor can not get higher profit by deviating to recommend either a serious treatment or a lab test. When the diagnosis is serious, doctor also will not deviate to a serious treatment which reuslts in rejection. ■

Proof of Lemma 4. Suppose there exists an equilibrium in which doctor only cheat on severity. In the equilibrium, no matter what the actual diagnosis is, doctor always reports the ture type of problem and a serious diagnosis. Assume that the equilibrium prices are $\{p_{LT}, p_{sT}, p_{mT}\}_{T\in\{A,B\}}$. In the equilibrium, kickback for each type of lab is k. The equilibrium acceptence rate for the two type of tests are equal:

$$\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} = \frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s}$$

Where $\hat{l}_{sA} = \hat{l}_{sB} = \hat{l}_{KK} = \frac{k+c_L}{\hat{\alpha}_{KK}} + p_s + \frac{(1-\hat{\alpha}_{KK})}{\hat{\alpha}_{KK}}(p_m - l_m)$, $\frac{1}{\hat{\alpha}_{KK}} = \frac{\omega}{\alpha} + \frac{1-\omega}{\tilde{\alpha}}$.

Doctor's expected profit is

$$\pi_{KK} = [\alpha(p_{sA} - c_s) + (1 - \alpha)(p_{mA} - c_m) + k]\frac{\bar{l}_s - \hat{l}_{KK}}{\bar{l}_s - \underline{l}_s}$$

Now we want to construct another equilibrium, in which doctor cheat on both severity and type of problem, and get higher profit than the equilibirum above.

Let $\hat{\alpha}_A^a = \frac{\alpha}{\alpha+(1-\alpha)(1+\theta)}$, $\hat{\alpha}_B^a = \frac{\alpha}{\alpha+(1-\alpha)(1-\theta)}$. In the equilibirum, doctor cheats on both type and severity. Unaware patients accept the recommendation for the lab test as long as

$$p_{LT}+\tilde{\alpha}p_{sT}+(1-\tilde{\alpha})p_{mT}\leq\tilde{\alpha}l_s+(1-\tilde{\alpha})l_m$$

Aware patients accepts the recommendation for the lab test as long as

$$p_{LT}+\hat{\alpha}_T^a p_{sT}+(1-\hat{\alpha}_T^a)p_{mT}\leq\hat{\alpha}_T^a l_s+(1-\hat{\alpha}_T^a)l_m$$

The average acceptence rate for each type of test is $\frac{\bar{l}_s-\hat{l}_{NK\_T}}{\bar{l}_s-\underline{l}_s}$, where

$$\hat{l}_{NK\_T}=\frac{k_T+c_L}{\hat{\alpha}_{NK\_T}}+p_{sT}+\frac{(1-\hat{\alpha}_{NK\_T})}{\hat{\alpha}_{NK\_T}}(p_{mT}-l_m)$$

$$\frac{1}{\hat{\alpha}_{NK\_T}}=\frac{\omega}{\hat{\alpha}_T^a}+\frac{1-\omega}{\tilde{\alpha}}$$

Then $\hat{\alpha}_{NK\_A}<\hat{\alpha}_{KK}<\hat{\alpha}_{NK\_B}$.

Then the profit of the doctor becomes

$$\begin{aligned}\pi_{NK} &= \frac{1}{2}[\alpha(p_{sA}-c_s)+(1-\alpha)(p_{mA}-c_m)+k-\varepsilon_1]\frac{\bar{l}_s-\hat{l}_{KK}}{\bar{l}_s-\underline{l}_s}\\ &+\frac{1}{2}[(1-\alpha)\theta(p_{mA}-c_m+k-\varepsilon_1)]\frac{\bar{l}_s-\hat{l}_{KK}}{\bar{l}_s-\underline{l}_s}\\ &+\frac{1}{2}[\alpha(p_{sB}-c_s)+(1-\alpha)(1-\theta)(p_{mB}-c_m)+(\alpha+(1-\alpha)(1-\theta))(k+\varepsilon_2)]\frac{\bar{l}_s-\hat{l}_{KK}}{\bar{l}_s-\underline{l}_s}\end{aligned}$$

If we keep all the prices same as in the orginal equilibrium, then

$$\begin{aligned}\pi_{NK}-\pi_{KK} &= \frac{1}{2}\frac{\hat{l}_{KK}-\hat{l}_{NK\_A}}{\bar{l}_s-\underline{l}_s}[\alpha(p_{sA}-c_s)+(1-\alpha)(p_{mA}-c_m)+k]\\ &+\frac{1}{2}\frac{\hat{l}_{KK}-\hat{l}_{NK\_A}}{\bar{l}_s-\underline{l}_s}[(1-\alpha)\theta(p_{mA}-c_m+k)]\\ &-\frac{1}{2}[\alpha(p_{sB}-c_s)+(1-\alpha)(1-\theta)(p_{mB}-c_m)+(\alpha+(1-\alpha)(1-\theta))k]\frac{\hat{l}_{NK\_B}-\hat{l}_{KK}}{\bar{l}_s-\underline{l}_s}\\ &= \alpha(p_s-c_s)[\frac{1}{2}\frac{\hat{l}_{KK}-\hat{l}_{NK\_A}}{\bar{l}_s-\underline{l}_s}-\frac{1}{2}\frac{\hat{l}_{NK\_B}-\hat{l}_{KK}}{\bar{l}_s-\underline{l}_s}]\\ &+\frac{1}{2}(p_{mA}-c_m)[(1-\alpha)(1+\theta)\frac{\hat{l}_{KK}-\hat{l}_{NK\_A}}{\bar{l}_s-\underline{l}_s}-(1-\alpha)(1-\theta)\frac{\hat{l}_{NK\_B}-\hat{l}_{KK}}{\bar{l}_s-\underline{l}_s}]\\ &+k[\frac{1}{2}[1+(1-\alpha)\theta]\frac{\hat{l}_{KK}-\hat{l}_{NK\_A}}{\bar{l}_s-\underline{l}_s}-\frac{1}{2}[1+(1-\alpha)\theta]\frac{\hat{l}_{NK\_B}-\hat{l}_{KK}}{\bar{l}_s-\underline{l}_s}]\end{aligned}$$

Since

$$
\begin{aligned}
(\hat{l}_{KK}-\hat{l}_{NK\_A})-(\hat{l}_{NK\_B}-\hat{l}_{KK}) &= 2\frac{k+c_L}{\hat{\alpha}_{KK}}-\frac{k+c_L}{\hat{\alpha}_{NK\_B}}-\frac{k+c_L}{\hat{\alpha}_{NK\_A}} \\
&= (k+c_L)[2(\frac{\omega}{\alpha}+\frac{1-\omega}{\tilde{\alpha}})-(\frac{\alpha+(1-\alpha)(1+\theta)}{\alpha}\omega+\frac{1-\omega}{\tilde{\alpha}})-(\frac{\omega}{\tilde{\alpha}}+ \\
&= 0
\end{aligned}
$$

$$\pi_{NK}-\pi_{KK}>0$$

To satisfy the incentive condition for the doctor, we can set $k_{NK\_A}=k+\varepsilon$, $k_{NK\_B}=k$, since profit function is continuous, there should be some small enough $\varepsilon>0$, such that $\pi_{NK}-\pi_{KK}>0$ still holds and IC conditions for the doctor below holds:

$$(p_{mA}-c_m+k+\varepsilon)]\frac{\bar{l}_s-\hat{l}_{NK\_A}}{\bar{l}_s-\underline{l}_s}>(p_{mB}-c_m+k)\frac{\bar{l}_s-\hat{l}_{NK\_B}}{\bar{l}_s-\underline{l}_s}$$

Since in the orginal equilibrium $(p_{mA}-c_m+k)\frac{\bar{l}_s-\hat{l}_{KK}}{\bar{l}_s-\underline{l}_s}\geq p_{mA}-c_m$, we have $(p_{mA}-c_m+k+\varepsilon)]\frac{\bar{l}_s-\hat{l}_{NK\_A}}{\bar{l}_s-\underline{l}_s}\geq p_{mA}-c_m$ is also satisfied here. ■

Proof of Proposition 3. Now we only need to solve the optimal problem maximizing doctors expected profit:

$$
\begin{aligned}
\pi^u_{NK} &= \frac{1}{2}[\alpha(p_{sA}-c_s)+(1-\alpha)(p_{mA}-c_m)+k_A]\frac{\bar{l}_s-l^u_{NK\_A}}{\bar{l}_s-\underline{l}_s} \\
&+\frac{1}{2}[(1-\alpha)\theta(p_{mA}-c_m+k_A)]\frac{\bar{l}_s-l^u_{NK\_A}}{\bar{l}_s-\underline{l}_s} \\
&+\frac{1}{2}[\alpha(p_{sB}-c_s)+(1-\alpha)(1-\theta)(p_{mB}-c_m)]\frac{\bar{l}_s-l^u_{NK\_B}}{\bar{l}_s-\underline{l}_s}
\end{aligned}
$$

$$c_s \leq p_{sT}\leq l^u_{NK\_T} \tag{13}$$

$$c_m \leq p_{mT}\leq l_m \tag{14}$$

$$\underline{l}_s \leq l_{sT}\leq \bar{l}_s \tag{15}$$

$$\frac{\bar{l}_s-l^u_{NK\_A}}{\bar{l}_s-\underline{l}_s}(k_T+p_{mA}-c_m)\geq p_{mA}-c_m \tag{16}$$

$$\begin{aligned} p_{sA} &= l^u_{NK\_A} + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mA}) - \frac{p_{LA}}{\tilde{\alpha}} \\ &= l^u_{NK\_A} + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mA}) - \frac{k_A + c_L}{\tilde{\alpha}} \\ p_{sB} &= l^u_{NK\_B} + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mB}) - \frac{p_{LB}}{\tilde{\alpha}} \\ &= l^u_{NK\_B} + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mB}) - \frac{k_A + c_L}{\tilde{\alpha}} \end{aligned}$$

$$\begin{aligned} \pi^u_{NK} &= \frac{1}{2}[\alpha(l^u_{NK\_A} - c_s) + \alpha\frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mA}) - \alpha\frac{k_A + c_L}{\tilde{\alpha}} + (1-\alpha)(p_{mA} - c_m) + k_A]\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s} \\ &+\frac{1}{2}[(1-\alpha)\theta(p_{mA} - c_m + k_A)]\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s} \\ &+\frac{1}{2}[\alpha(l^u_{NK\_B} - c_s) + \alpha\frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mB}) - \alpha\frac{c_L}{\tilde{\alpha}} + (1-\alpha)(1-\theta)(p_{mB} - c_m)]\frac{\bar{l}_s - l^u_{NK\_B}}{\bar{l}_s - \underline{l}_s} \end{aligned}$$

The doctor can separately set prices for different types of problem.

For type A:

$$\begin{aligned} \pi^u_{NK\_A} &= \frac{1}{2}[\alpha(l^u_{NK\_A} - c_s) + \alpha\frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mA}) \\ &-\alpha\frac{k_A + c_L}{\tilde{\alpha}} + (1-\alpha)(p_{mA} - c_m) + k_A]\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s} \\ &+\frac{1}{2}[(1-\alpha)\theta(p_{mA} - c_m + k_A)]\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s} \end{aligned}$$

$$\begin{aligned} L^A &= \frac{1}{2}[\alpha(l_{sA} - c_s) + \alpha\frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mA}) - \alpha\frac{k_A + c_L}{\tilde{\alpha}} + (1-\alpha)(p_{mA} - c_m) + k_A]\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s} \\ &+\frac{1}{2}[(1-\alpha)\theta(p_{mA} - c_m + k_A)]\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s} + \lambda[l^u_{NK\_A} + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mA}) - \frac{k_A + c_L}{\tilde{\alpha}} - c_s] \end{aligned}$$

$$\frac{\partial L}{\partial k_A} = \frac{1}{2}\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s}[1 - \frac{\alpha}{\tilde{\alpha}} + (1-\alpha)\theta] - \frac{\lambda}{\tilde{\alpha}} \leq 0 \qquad (\,= 0 \text{ if } k_s > 0\,)$$

$$\begin{aligned} \lambda[l^u_{NK\_A} + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mA}) - \frac{k_A + c_L}{\tilde{\alpha}} - c_s] &= 0 \\ \lambda &\geq 0 \\ l^u_{NK\_A} + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mA}) - \frac{k_A + c_L}{\tilde{\alpha}} - c_s &\geq 0 \end{aligned}$$

$$\frac{1}{2}\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s}[1 - \frac{\alpha}{\tilde{\alpha}} + (1-\alpha)\theta] - \frac{\lambda}{\tilde{\alpha}} \leq 0$$

$$\Longrightarrow \frac{\lambda}{\tilde{\alpha}} \geq \frac{1}{2}\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s}[1 - \frac{\alpha}{\tilde{\alpha}} + (1-\alpha)\theta] > 0$$
$$\Longrightarrow l_{sA} + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mA}) - \frac{k_A + c_L}{\tilde{\alpha}} - c_s = 0$$
$$\Longrightarrow k_A = \tilde{\alpha}(l^u_{NK\_A} - c_s) + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mA}) - c_L > 0$$
$$\Longrightarrow \frac{1}{2}\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s}[1 - \frac{\alpha}{\tilde{\alpha}} + (1-\alpha)\theta] - \frac{\lambda}{\tilde{\alpha}} = 0$$
$$\Longrightarrow \lambda = \frac{1}{2}\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s}[\tilde{\alpha} - \alpha + \tilde{\alpha}(1-\alpha)\theta]$$

$$\begin{aligned}
\frac{\partial L^A}{\partial p_{mA}} &= \frac{1}{2}\frac{\bar{l}_s - l^u_{NK_A}}{\bar{l}_s - \underline{l}_s}[-\alpha\frac{(1-\tilde{\alpha})}{\tilde{\alpha}} + (1-\alpha) + (1-\alpha)\theta] - \lambda\frac{(1-\tilde{\alpha})}{\tilde{\alpha}} \\
&= \frac{1}{2}\frac{\bar{l}_s - l^u_{NK_A}}{\bar{l}_s - \underline{l}_s}[-\alpha\frac{(1-\tilde{\alpha})}{\tilde{\alpha}} + (1-\alpha) + (1-\alpha)\theta - \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(\tilde{\alpha} - \alpha + \tilde{\alpha}(1-\alpha)\theta)] \\
&= \frac{1}{2}\frac{\bar{l}_s - l^u_{NK_A}}{\bar{l}_s - \underline{l}_s}[\tilde{\alpha} - \alpha + (1-\alpha)\theta] > 0
\end{aligned}$$

Then $p_{mA} = l_m$.

$$\Longrightarrow k_A = \tilde{\alpha}(l^u_{NK\_A} - c_s) - c_L$$

$$\begin{aligned}
\pi^u_{NK_A} &= \frac{1}{2}[\alpha(l^u_{NK_A} - c_s) - \alpha\frac{k_A + c_L}{\tilde{\alpha}} + (1-\alpha)(l_m - c_m) + k_A]\frac{\bar{l}_s - l^u_{NK_A}}{\bar{l}_s - \underline{l}_s} \\
&\quad + \frac{1}{2}[(1-\alpha)\theta(l_m - c_m + k_A)]\frac{\bar{l}_s - l^u_{NK_A}}{\bar{l}_s - \underline{l}_s}
\end{aligned}$$

$$\begin{aligned}
L^A &= \frac{1}{2}[\alpha(l^u_{NK_A} - c_s) - \alpha\frac{k_A + c_L}{\tilde{\alpha}} + (1-\alpha)(l_m - c_m) + k_A]\frac{\bar{l}_s - l^u_{NK_A}}{\bar{l}_s - \underline{l}_s} \\
&\quad + \frac{1}{2}[(1-\alpha)\theta(l_m - c_m + k_A)]\frac{\bar{l}_s - l^u_{NK_A}}{\bar{l}_s - \underline{l}_s} \\
&\quad + \lambda[l^u_{NK_A} - \frac{k_A + c_L}{\tilde{\alpha}} - c_s]
\end{aligned}$$

$$\frac{\partial L^A}{\partial l^u_{NK\_A}} = \frac{1}{2}\alpha\frac{\bar{l}_s - l^u_{NK\_A}}{\bar{l}_s - \underline{l}_s} - \frac{1}{\bar{l}_s - \underline{l}_s}\frac{1}{2}[\alpha(l^u_{NK\_A} - c_s) - \alpha\frac{k_A + c_L}{\tilde{\alpha}} + (1-\alpha)(l_m - c_m) + k_A]$$
$$-\frac{1}{2}[(1-\alpha)\theta(l_m - c_m + k_A)]\frac{1}{\bar{l}_s - \underline{l}_s} + \lambda \le 0 \quad (=0 \text{ if } l_{sA} > \underline{l}_s)$$

$$\begin{aligned}
\frac{\partial L^A}{\partial l^u_{NK_A}} &= \frac{1}{2}\alpha\frac{\bar{l}_s - l^u_{NK_A}}{\bar{l}_s - \underline{l}_s} - \frac{1}{\bar{l}_s - \underline{l}_s}\frac{1}{2}[\alpha(l^u_{NK_A} - c_s) - \alpha\frac{k_A + c_L}{\tilde{\alpha}} + (1-\alpha)(l_m - c_m) + k_A] \\
&\quad -\frac{1}{2}[(1-\alpha)\theta(l_m - c_m + k_A)]\frac{1}{\bar{l}_s - \underline{l}_s} + \frac{1}{2}\frac{\bar{l}_s - l^u_{NK_A}}{\bar{l}_s - \underline{l}_s}(\tilde{\alpha} - \alpha + \tilde{\alpha}(1-\alpha)\theta) \\
&= \frac{1}{2}\frac{1}{\bar{l}_s - \underline{l}_s}[\tilde{\alpha}(\bar{l}_s - l^u_{NK_A}) - \alpha(l^u_{NK_A} - c_s) + \alpha\frac{k_A + c_L}{\tilde{\alpha}} - (1-\alpha)(l_m - c_m) - k_A \\
&\quad -(1-\alpha)\theta(l_m - c_m + k_A) + \tilde{\alpha}(1-\alpha)\theta(\bar{l}_s - l_{sA})] \\
&= \frac{1}{2}\frac{1}{\bar{l}_s - \underline{l}_s}[\tilde{\alpha}(\bar{l}_s - l^u_{NK_A}) - \alpha(l^u_{NK_A} - c_s) + \alpha(l_{sA} - c_s) \\
&\quad -(1-\alpha)(l_m - c_m) - \tilde{\alpha}(l^u_{NK_A} - c_s) + c_L \\
&\quad -(1-\alpha)\theta(l_m - c_m + \tilde{\alpha}(l^u_{NK_A} - c_s) - c_L) + \tilde{\alpha}(1-\alpha)\theta(\bar{l}_s - l^u_{NK_A})] \\
&= \frac{1}{2}\frac{1}{\bar{l}_s - \underline{l}_s}\{\tilde{\alpha}[1 + (1-\alpha)\theta](\bar{l}_s - 2l^u_{NK_A} + c_s + \frac{c_L}{\tilde{\alpha}}) - (1-\alpha)(1+\theta)(l_m - c_m)\}
\end{aligned}$$

$$\frac{\partial L}{\partial l^u_{NK\_A}} = 0 \Leftrightarrow \tilde{\alpha}[1 + (1-\alpha)\theta](\bar{l}_s - 2l^u_{NK\_A} + c_s + \frac{c_L}{\tilde{\alpha}}) - (1-\alpha)(1+\theta)(l_m - c_m) = 0$$

$$\begin{aligned}
l^u_{NK_A} &= \frac{\tilde{\alpha}[1 + (1-\alpha)\theta](\bar{l}_s + c_s + c_L) - (1-\alpha)(1+\theta)(l_m - c_m)}{2\tilde{\alpha}[1 + (1-\alpha)\theta]} \\
&= \bar{l}_s - \frac{\tilde{\alpha}(\bar{l}_s - c_s) + \frac{(1-\alpha)(1+\theta)}{1+(1-\alpha)\theta}(l_m - c_m) - c_L}{2\tilde{\alpha}}
\end{aligned}$$

this interior solution requires $\frac{\tilde{\alpha}[1+(1-\alpha)\theta](\bar{l}_s+c_s+c_L)-(1-\alpha)(1+\theta)(l_m-c_m)}{2\tilde{\alpha}[1+(1-\alpha)\theta]} > \underline{l}_s$.

In this case, $k_A = \tilde{\alpha}(l^u_{NK\_A} - c_s) - c_L = \frac{\tilde{\alpha}(\bar{l}_s-c_s)+\frac{(1-\alpha)(1+\theta)}{(1+(1-\alpha)\theta}(l_m-c_m)+c_L}{2}$.

Check constraint (16):

$$\begin{aligned}
\frac{\bar{l}_s - l^u_{NK_A}}{\bar{l}_s - \underline{l}_s}(k_A + l_m - c_m) &\ge l_m - c_m \\
\tilde{\alpha}(\bar{l}_s - l^u_{NK_A})(k_A + l_m - c_m) &\ge \tilde{\alpha}(l_m - c_m)(\bar{l}_s - \underline{l}_s)
\end{aligned}$$

since $\tilde{\alpha}(\bar{l}_s - l^u_{NK\_A}) < \tilde{\alpha}(\bar{l}_s - \underline{l}_s)$, then $k_A + l_m - c_m > l_m - c_m$ and $\tilde{\alpha}(\bar{l}_s - l_{sA}) + k_A + l_m - c_m = \tilde{\alpha}(\bar{l}_s - c_s) + l_m - c_m - c_L > (l_m - c_m) + \tilde{\alpha}(\bar{l}_s - \underline{l}_s)$, then we have (16) holds.

When $\frac{\tilde{\alpha}[1+(1-\alpha)\theta](\bar{l}_s+c_s+c_L)-(1-\alpha)(1+\theta)(l_m-c_m)}{2\tilde{\alpha}[1+(1-\alpha)\theta]} \leq \underline{l}_s$, the optimal solution is $l_{sA} = \underline{l}_s$ and $k_A = \tilde{\alpha}(\underline{l}_s - c_s) - c_L$. (16) automatically holds in this case.

Now we solve pricing problem for type B.

$$\pi^u_{NK\_B} = \frac{1}{2}[\alpha(l^u_{NK\_B}-c_s)+\alpha\frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m-p_{mB})-\alpha\frac{k_B+c_L}{\tilde{\alpha}}+(1-\alpha)(1-\theta)(p_{mB}-c_m)+\alpha\frac{k_B}{\tilde{\alpha}}]\frac{\bar{l}_s-l^u_{NK\_B}}{\bar{l}_s-\underline{l}_s}$$

$$p_{sB} = l^u_{NK\_B} + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(l_m - p_{mB}) - \frac{k_B+c_L}{\tilde{\alpha}}$$

$p_{sB} > c_s$ in the optimal solution, since by assumption $c_s < \underline{l}_s - \frac{c_L}{\tilde{\alpha}}$

Remark: $\alpha\frac{(1-\tilde{\alpha})}{\tilde{\alpha}} = (1-\alpha)(1-\theta)$

$$\pi^u_{NK\_B} = \frac{1}{2}[\alpha(l^u_{NK\_B}-c_s) - \alpha\frac{c_L}{\tilde{\alpha}} + (1-\alpha)(1-\theta)(l_m-c_m)]\frac{\bar{l}_s-l^u_{NK\_B}}{\bar{l}_s-\underline{l}_s}$$

Then price of minor treatment for problem B will not affect the profit of doctor.

$$\begin{aligned}
\frac{\partial\pi^u_{NK_B}}{\partial l^u_{NK_B}} &= \frac{1}{2}\alpha\frac{\bar{l}_s-l^u_{NK_B}}{\bar{l}_s-\underline{l}_s} - \frac{1}{2}[\alpha(l^u_{NK_B}-c_s)+(1-\alpha)(1-\theta)(l_m-c_m)-\alpha\frac{c_L}{\tilde{\alpha}}]\frac{1}{\bar{l}_s-\underline{l}_s} \\
&= \frac{1}{2}\frac{1}{\bar{l}_s-\underline{l}_s}[\alpha\bar{l}_s - 2\alpha l^u_{NK_B} + \alpha c_s - (1-\alpha)(1-\theta)(l_m-c_m) + \alpha\frac{c_L}{\tilde{\alpha}}]
\end{aligned}$$

If $l^u_{NK\_B} > \underline{l}_s \Longrightarrow \frac{\partial\pi^u_{NK\_B}}{\partial l^u_{NK\_B}} = 0 \Longrightarrow \alpha\bar{l}_s - 2\alpha l^u_{NK\_B} + \alpha c_s - (1-\alpha)(1-\theta)(l_m-c_m) + \alpha\frac{c_L}{\tilde{\alpha}} = 0$

$$\begin{aligned}
l^u_{NK_B} &= \frac{\alpha\bar{l}_s + \alpha c_s - (1-\alpha)(1-\theta)(l_m-c_m) + \alpha\frac{c_L}{\tilde{\alpha}}c_L}{2\alpha} \\
&= \bar{l}_s - \frac{\alpha(\bar{l}_s-c_s) + (1-\alpha)(1-\theta)(l_m-c_m) - \alpha\frac{c_L}{\tilde{\alpha}}c_L}{2\alpha} \\
&= \bar{l}_s - \frac{\tilde{\alpha}(\bar{l}_s-c_s) + (1-\tilde{\alpha})(l_m-c_m) - c_l}{2\tilde{\alpha}}
\end{aligned}$$

■

Proof of Corollary 1. It is obviously that when patients are unaware of kickback, the result will be the same as in the case of no kickback.

$$l^u_{NN} = \bar{l}_s - \frac{\tilde{\alpha}(\bar{l}_s-c_s) + (1-\tilde{\alpha})(l_m-c_m) - c_L}{2\tilde{\alpha}}$$

Kickback will increase the accept rate of lab test.

$$l^u_{NK\_A} = \bar{l}_s - \frac{\tilde{\alpha}(\bar{l}_s-c_s) + \frac{(1-\alpha)(1+\theta)}{1+(1-\alpha)\theta}(l_m-c_m) - c_L}{2\tilde{\alpha}}$$

$$l^u_{NK\_A} < l^u_{NN} = l^u_{NK\_B}$$

■

Proof of Lemma 5. If the doctor decides to accept kickback from both labs, aware consumers will undertand the doctor's strategy. They will accept a lab test if and only if

$$p_L + \alpha p_s + (1-\alpha)p_m \le \alpha l_s + (1-\alpha)l_m$$

The expected profit of doctor is given by:

$$\begin{aligned}
\pi^a_{KK} &= \frac{1}{2}[\alpha(p_{sA}-c_s) + (1-\alpha)(p_{mA}-c_m) + k_A]\frac{\bar{l}_s - l^a_{KK_A}}{\bar{l}_s - \underline{l}_s} \\
&\quad + \frac{1}{2}[\alpha(p_{sB}-c_s) + (1-\alpha)(p_{mB}-c_m) + k_B]\frac{\bar{l}_s - l^a_{KK_B}}{\bar{l}_s - \underline{l}_s} \\
&= \frac{1}{2}[\alpha l^a_{KK_A} - \alpha c_s + (1-\alpha)(l_m - c_m) - c_{LA}]\frac{\bar{l}_s - l^a_{KK_A}}{\bar{l}_s - \underline{l}_s} \\
&\quad + \frac{1}{2}[\alpha l^a_{KK_B} - \alpha c_s + (1-\alpha)(l_m - c_m) - c_{LB}]\frac{\bar{l}_s - l^a_{KK_B}}{\bar{l}_s - \underline{l}_s}
\end{aligned}$$

Since two labs are symmetric, the optimal pricing for each type of problem should be the same. As a price taker, each lab sets price at $p_{LT} = k_T + c_L$, where $T \in \{A, B\}$. Then it is equivalent to maximize:

$$\pi^a_{KK} = [\alpha l^a_{KK} + (1-\alpha)(l_m - c_m) - c_{LA} - \alpha c_s]\frac{\bar{l}_s - l^a_{KK}}{\bar{l}_s - \underline{l}_s}$$

$$\begin{aligned}
L^a_{KK} &= [\alpha l^a_{KK} + (1-\alpha)(l_m - c_m) - c_{LA} - \alpha c_s]\frac{\bar{l}_s - l^a_{KK}}{\bar{l}_s - \underline{l}_s} \\
&\quad + \lambda[l^a_{KK} + \frac{(1-\hat{\alpha})}{\hat{\alpha}}(l_m - p_m) - \frac{k_A + c_L}{\hat{\alpha}} - c_s]
\end{aligned}$$

It is easy to see that the kickback will not affect the profit of doctor, in this case,

FOCs:

$$\begin{aligned}
\frac{\partial L^a_{KK}}{\partial p_m} &= -\lambda\frac{(1-\alpha)}{\alpha} \le 0 \\
\frac{\partial L^a_{KK}}{\partial k_A} &= -\frac{\lambda}{\alpha} \le 0
\end{aligned}$$

$$\frac{\partial L^a_{KK}}{\partial l^a_{KK}} = [\alpha l^a_{KK} + (1-\alpha)(l_m - c_m) - c_{LA} - \alpha c_s]\frac{-1}{\bar{l}_s - \underline{l}_s} + \alpha\frac{\bar{l}_s - l^a_{KK}}{\bar{l}_s - \underline{l}_s} + \lambda \le 0 \qquad (=0 \text{ if } l_{sA} > \underline{l}_s)$$

If $\lambda > 0$, Then $k_A = 0$ $p_m = c_m$ and $l^a_{KK} + \frac{(1-\hat{\alpha})}{\hat{\alpha}}(l_m - p_m) - \frac{k_A+c_L}{\hat{\alpha}} - c_s = 0$, contradiction to our assumption.

Then $\lambda = 0$

$$\begin{aligned}
\frac{\partial L^a_{KK}}{\partial l^a_{KK}} &= [\alpha(l^a_{KK} - c_s) + (1-\alpha)(l_m - c_m) - c_L]\frac{-1}{\bar{l}_s - \underline{l}_s} + \alpha\frac{\bar{l}_s - l^a_{KK}}{\bar{l}_s - \underline{l}_s} \leq 0 \\
\frac{\partial L^a_{KK}}{\partial l^a_{KK}} &= \frac{1}{\bar{l}_s - \underline{l}_s}[\alpha(\bar{l}_s - l^a_{KK}) - \alpha(l^a_{KK} - c_s) - (1-\alpha)(l_m - c_m) + c_L] \\
&= \frac{1}{\bar{l}_s - \underline{l}_s}[\alpha\bar{l}_s - 2\alpha l^a_{KK} + \alpha c_s - (1-\alpha)(l_m - c_m) + c_L]
\end{aligned}$$

If $\frac{\alpha\bar{l}_s+\alpha c_s-(1-\alpha)(l_m-c_m)+c_L}{2\alpha} < \underline{l}_s \Longrightarrow \frac{\partial L^a_{KK}}{\partial l^a_{KK}} < 0$, then $l^a_{KK} = \underline{l}_s$, in this case, $k_A = \hat{\alpha}(\underline{l}_s - c_s) - c_L$

If $\frac{\alpha\bar{l}_s+\alpha c_s-(1-\alpha)(l_m-c_m)+c_L}{2\alpha} \geq \underline{l}_s \Longrightarrow \frac{\partial L^a_{KK}}{\partial l^a_{KK}} = 0$, we get interior solution:

$$\begin{aligned}
l^a_{KK} &= \frac{\alpha\bar{l}_s + \alpha c_s - (1-\alpha)(l_m - c_m) + c_L}{2\alpha} \\
&= \bar{l}_s - \frac{\alpha(\bar{l}_s - c_s) + (1-\alpha)(l_m - c_m) - c_L}{2\alpha}
\end{aligned}$$

The optimal price of type B problem is the same.

$$\pi^a_{KK} = \begin{cases} \alpha\frac{(\bar{l}_s - l^a_{KK})^2}{\bar{l}_s - \underline{l}_s} & \text{if } \ l^a_{KK} > \underline{l}_s \\ \alpha(\underline{l}_s - c_s) + (1-\alpha)(l_m - c_m) - c_L & \text{if } \ l^a_{KK} = \underline{l}_s \end{cases}$$

Compared with the situation doctor decides to not charge any kickback,

$$l^a_{NN} = l^n = \bar{l}_s - \frac{\tilde{\alpha}(\bar{l}_s - c_s) + (1-\tilde{\alpha})(l_m - c_m) - c_L}{2\tilde{\alpha}}$$

$$\pi^a_{NN} = \pi^n = \begin{cases} (1-\alpha)\,\theta(l_m - c_m) + \frac{\alpha(\bar{l}_s - l^{nk})^2}{\bar{l}_s - \underline{l}_s} \text{ if } l^a_{NN} > \underline{l}_s \\ \alpha(\underline{l}_s - c_s) + (1-\alpha)(l_m - c_m) - \frac{\alpha}{\tilde{\alpha}}c_L \end{cases}$$

1) When $l^n = l^a_{KK} = \underline{l}_s$

$$\pi^a_{NN} - \pi^a_{KK} = (1-\alpha)\,\theta c_L > 0$$

2) when $l^n > l^a_{KK} = \underline{l}_s$

$$\begin{aligned}
\pi^a_{NN} - \pi^a_{KK} &= \frac{\alpha(\bar{l}_s - l^n)^2}{\bar{l}_s - \underline{l}_s} + (1-\alpha)\,\theta(l_m - c_m) - \alpha(\underline{l}_s - c_s) - (1-\alpha)(l_m - c_m) + c_L \\
&= \frac{\alpha(l^n - \underline{l}_s)^2}{\bar{l}_s - \underline{l}_s} + (1-\alpha)\theta c_L > 0
\end{aligned}$$

3) when $l^n > l^a_{KK} > \underline{l}_s$

$$\begin{aligned}
\pi^a_{KK} - \pi^a_{NN} &= \frac{\alpha(\bar{l}_s - l^a_{KK})^2}{\bar{l}_s - \underline{l}_s} - \frac{\alpha(\bar{l}_s - l^n)^2}{\bar{l}_s - \underline{l}_s} - (1-\alpha)\,\theta(l_m - c_m) \\
&= \frac{\alpha(l^n - l^a_{KK})(2\bar{l}_s - l^a_{KK} - l^n)}{\bar{l}_s - \underline{l}_s} - 2\alpha(l^n - l^a_{KK}) - \theta\,(1-\alpha)\,c_L \\
&= \frac{\alpha(l^n - l^a_{KK})}{\bar{l}_s - \underline{l}_s}[2\bar{l}_s - \widehat{l}^{ak}_s - \widehat{l}^{nk}_s - 2(\bar{l}_s - \underline{l}_s)] - \theta\,(1-\alpha)\,c_L \\
&= \frac{\alpha(l^n - l^a_{KK})}{\bar{l}_s - \underline{l}_s}(2\underline{l}_s - l^a_{KK} - l^n) - \theta\,(1-\alpha)\,c_L < 0
\end{aligned}$$

For the asymmetric case:

If $l^a_{NK\_B} = l^n > l^a_{NK\_A} > \underline{l}_s$

$$\begin{aligned}
\pi^a_{NK} - \pi^a_{NN} &= \frac{1}{2}\alpha\frac{(\bar{l}_s - l^a_{NK_A})^2}{\bar{l}_s - \underline{l}_s} + \frac{1}{2}\alpha\frac{(\bar{l}_s - l^a_{NK_B})^2}{\bar{l}_s - \underline{l}_s} \\
&\quad - \frac{\alpha(\bar{l}_s - l^n)^2}{\bar{l}_s - \underline{l}_s} - (1-\alpha)\,\theta(l_m - c_m) \\
&= \frac{1}{2}\alpha\frac{(\bar{l}_s - l^a_{NK_A})^2}{\bar{l}_s - \underline{l}_s} - \frac{1}{2}\frac{\alpha(\bar{l}_s - l^n)^2}{\bar{l}_s - \underline{l}_s} - (1-\alpha)\,\theta(l_m - c_m) \\
&= \frac{1}{2}\alpha\frac{(l^n - l^a_{NK_A})}{\bar{l}_s - \underline{l}_s}(2\bar{l}_s - l^a_{NK_A} - l^n) - (1-\alpha)\,\theta(l_m - c_m) \\
&= \frac{1}{2}\alpha\frac{2\bar{l}_s - l^a_{NK_A} - l^n}{\bar{l}_s - \underline{l}_s}(l_m - c_m - c_L)(\frac{1}{2\hat{\alpha}_A} - \frac{1}{2\tilde{\alpha}}) - (1-\alpha)\,\theta(l_m - c_m) \\
&= \frac{1}{2}\frac{2\bar{l}_s - l^a_{NK_A} - l^n}{\bar{l}_s - \underline{l}_s}(l_m - c_m - c_L)\theta(1-\alpha) - (1-\alpha)\,\theta(l_m - c_m) < 0
\end{aligned}$$

if $l^a_{NK\_B} = l^n > \underline{l}_s = l^a_{NK\_A}$

$$\begin{aligned}
\pi^a_{NK} - \pi^a_{NN} &= \frac{1}{2}[\alpha + (1-\alpha)(1+\theta)][\hat{\alpha}_A(\underline{l}_s - c_s) + (1-\hat{\alpha}_A)(l_m - c_m) - c_L] \\
&\quad + \frac{1}{2}\alpha\frac{(\bar{l}_s - l^a_{NK_B})^2}{\bar{l}_s - \underline{l}_s} - \frac{\alpha(\bar{l}_s - l^n)^2}{\bar{l}_s - \underline{l}_s} - (1-\alpha)\,\theta(l_m - c_m) \\
&< \frac{1}{2}\alpha\frac{(\bar{l}_s - \frac{\hat{\alpha}_A(\bar{l}_s + c_s) - (1-\hat{\alpha}_A)(l_m - c_m) + c_L}{2\hat{\alpha}_A})^2}{\bar{l}_s - \underline{l}_s} - \frac{1}{2}\frac{\alpha(\bar{l}_s - l^n)^2}{\bar{l}_s - \underline{l}_s} - (1-\alpha)\,\theta(l_m - c_m) < 0
\end{aligned}$$

if $l^a_{NK\_B} = l^n = \underline{l}_s = l^a_{NK\_A}$, it is obvious $\pi^a_{NK} - \pi^a_{NN} < 0$ as more cost is wasted on the lab test.

Then we can show that accept kickback from both labs will be non-profitable compared with no kickback case, then doctor will not charge kickback from both labs. ■

Proof of proposition ??. In the optimal equilibrium that there is mixed population of aware and unaware patients, doctor's expected profit is given by:

$$\pi^a = \frac{1}{2}\hat{\alpha}_A[(1-\alpha)\theta+1]\frac{(\bar{l}_s - \hat{l}_{sA})^2}{\bar{l}_s - \underline{l}_s} + \frac{1}{2}\alpha\frac{(\bar{l}_s - \hat{l}_{sB})^2}{\bar{l}_s - \underline{l}_s}$$

$$\hat{l}_{sA} = \frac{\hat{\alpha}_A(\bar{l}_s + c_s) - \frac{(1-\alpha)(1+\theta)}{1+(1-\alpha)\theta}(l_m - c_m) + c_L}{2\hat{\alpha}_A}$$

$$\hat{l}_{sB} = \frac{\tilde{\alpha}(\bar{l}_s + c_s) - (1-\tilde{\alpha})(l_m - c_m) + c_L}{2\tilde{\alpha}}$$

$$\hat{\alpha}_A = \frac{\alpha}{1+(2\omega-1)(1-\alpha)\theta}$$

$$\begin{aligned}
\frac{\partial \pi^a}{\partial \omega} &= \frac{\partial \pi^a}{\partial \hat{\alpha}_A}\frac{\partial \hat{\alpha}_A}{\partial \omega} \\
&= \frac{\partial \hat{\alpha}_A}{\partial \omega}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}\frac{[(1-\alpha)\theta+1](\bar{l}_s - \hat{l}_{sA}) - (1-\alpha)(1+\theta)(l_m - c_m) + [1+(1-\alpha)\theta]c_L}{2\hat{\alpha}_A} \\
&< 0
\end{aligned}$$

Since π^a is continuous, suppose $\pi^a|_{\omega=0} > 0$, then there exsit some $\hat{\omega} \in [0,1]$, such that when $\omega \geq \hat{\omega}$, doctor won't charge kickback. ■

9 Appendix B: Optimal equilibrium with a mixed population of unaware and aware patients

We focus on the pure strategy of each player. In the equilibirum, doctor cheats on both type and severity:

Unaware patients: always hold belief

$$\hat{\alpha}^u_A = \hat{\alpha}^u_B = \tilde{\alpha} = \frac{\alpha}{\alpha + (1-\alpha)(1+\theta)}$$

when he is recommended a lab test.

Aware patients: $\{\hat{\alpha}^a_A, \hat{\alpha}^a_B, \hat{k}_A, \hat{k}_B\}$ based on the prices she observed.

Aware patients know there is posibility that doctor charges kickback and she also know the fraction of aware patients. She will accept the recommendation for lab test if and only if:

$$p_{LA} + \hat{\alpha}^a_A p_{sA} + (1 - \hat{\alpha}^a_A) p_{mA} \leq \hat{\alpha}^a_A l_s + (1 - \hat{\alpha}^a_A) l_m$$

$$p_{LB} + \hat{\alpha}^a_B p_{sB} + (1 - \hat{\alpha}^a_B) p_{mB} \leq \hat{\alpha}^a_B l_s + (1 - \hat{\alpha}^a_B) l_m$$

Then she can evaluate the expected acceptence rate among all patients:

$$\begin{aligned}
\hat{l}_{sA} &= \omega \hat{l}^a_{sA} + (1-\omega)\hat{l}^u_{sA} \\
&= \omega[\frac{p_{LA}}{\hat{\alpha}^a_A} + p_{sA} + \frac{(1-\hat{\alpha}^a_A)}{\hat{\alpha}^a_A}(p_{mA} - l_m)] \\
&\quad +(1-\omega)[\frac{p_L}{\tilde{\alpha}} + p_{sA} + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(p_{mA} - l_m)] \\
&= \frac{p_{LA}}{\hat{\alpha}_A} + p_{sA} + \frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A}(p_{mA} - l_m)
\end{aligned}$$

Where $\frac{1}{\hat{\alpha}_A} = \frac{\omega}{\hat{\alpha}^a_A} + \frac{1-\omega}{\tilde{\alpha}}$

$$\begin{aligned}
\hat{l}_{sB} &= \omega \hat{l}^a_{sB} + (1-\omega)\hat{l}^u_{sB} \\
&= \omega[\frac{p_{LB}}{\hat{\alpha}^B_B} + p_{sB} + \frac{(1-\hat{\alpha}^a_B)}{\hat{\alpha}^B_B}(p_{mB} - l_m)] \\
&\quad +(1-\omega)[\frac{p_L}{\tilde{\alpha}} + p_{sB} + \frac{(1-\tilde{\alpha})}{\tilde{\alpha}}(p_{mB} - l_m)] \\
&= \frac{p_{LB}}{\hat{\alpha}_B} + p_{sB} + \frac{(1-\hat{\alpha}_B)}{\hat{\alpha}_B}(p_{mB} - l_m)
\end{aligned}$$

Where $\hat{\alpha}_B = \hat{\alpha}^B_B = \tilde{\alpha} = \frac{\alpha}{\alpha+(1-\alpha)(1+\theta)}$

Aware patients beliefs should conisidence with the following rule:

if $\hat{\alpha}^a_A = \frac{\alpha}{1+(1-\alpha)\theta}, \hat{\alpha}^a_B = \frac{\alpha}{1-(1-\alpha)\theta}$

$$\begin{aligned}
\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(\hat{k}_A + p_{mA} - c_m) &> \frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s}(\hat{k}_B + p_{mB} - c_m) \\
\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(\hat{k}_A + p_{mA} - c_m) &\geq l_m - c_m
\end{aligned}$$

if $\hat{\alpha}^a_A = \alpha = \hat{\alpha}^a_B$, then

$$\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(\hat{k}_A + p_{mA} - c_m) = \frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s}(\hat{k}_B + p_{mB} - c_m) \geq l_m - c_m$$

If $\hat{\alpha}_A^a = \frac{\alpha}{1-(1-\alpha)\theta} = \hat{\alpha}_B^a$, then

$$\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}(\hat{k}_A + p_{mA} - c_m) = \frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s}(\hat{k}_B + p_{mB} - c_m) \leq l_m - c_m$$

Doctor:

Given the beliefs of patients, after diagnosis, doctor chooses his recommendation strategy.

As long as

$$\frac{\bar{l}_s - \frac{p_L}{\hat{\alpha}_A} - p_{sA} - \frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A}(p_{mA} - l_m)}{\bar{l}_s - \underline{l}_s}(k_A + p_{mA} - c_m) > \frac{\bar{l}_s - \frac{p_L}{\hat{\alpha}_B} - p_{sB} - \frac{(1-\hat{\alpha}_B)}{\hat{\alpha}_B}(p_{mB} - l_m)}{\bar{l}_s - \underline{l}_s}(k_B + p_{mB} - c_m) \quad (17)$$

Doctor will cheat on both type and severity.

$$\begin{aligned}\pi \;=\; & \frac{1}{2}[\alpha(p_{sA} - c_s) + (1-\alpha)(p_{mA} - c_m) + k_A]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} \\ & + \frac{1}{2}[(1-\alpha)\theta(p_{mA} - c_m + k_A)]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} \\ & + \frac{1}{2}[\alpha(p_{sB} - c_s) + (1-\alpha)(1-\theta)(p_{mB} - c_m) + (\alpha + (1-\alpha)(1-\theta))k_B]\frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s}\end{aligned}$$

Fix any prices of lab test and treatments, in the sub-recommendation game, given patients' beliefs discribed above, there is no incentive for the doctor to deviate in his recommendation strategy. Aware patients' belief is fulfilled in the equilibrium. What left is to solve the optimal prices at the beginning of the game.

$$\begin{aligned}\pi \;=\; & \frac{1}{2}[\alpha(p_{sA} - c_s) + (1-\alpha)(p_{mA} - c_m) + k_A]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} \\ & + \frac{1}{2}[(1-\alpha)\theta(p_{mA} - c_m + k_A)]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} \\ & + \frac{1}{2}[\alpha(p_{sB} - c_s) + (1-\alpha)(1-\theta)(p_{mB} - c_m) + (\alpha + (1-\alpha)(1-\theta))k_B]\frac{\bar{l}_s - l^a_{NK\_B}}{\bar{l}_s - \underline{l}_s}\end{aligned}$$

For simplicity, we assume an **interior solution** at first. Then we have

$$
\begin{aligned}
p_{sA} &= \hat{l}_{sA} + \frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A}(l_m - p_{mA}) - \frac{p_{LA}}{\hat{\alpha}_A} \\
&= \hat{l}_{sA} + \frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A}(l_m - p_{mA}) - \frac{k_A + c_L}{\hat{\alpha}} \\
p_{sB} &= \hat{l}_{sB} + \frac{(1-\hat{\alpha}_B)}{\hat{\alpha}_B}(l_m - p_{mB}) - \frac{p_{LB}}{\hat{\alpha}_B} \\
&= \hat{l}_{sB} + \frac{(1-\hat{\alpha}_B)}{\hat{\alpha}_B}(l_m - p_{mB}) - \frac{k_B + c_L}{\hat{\alpha}_B}
\end{aligned}
$$

$$
\begin{aligned}
\pi &= \frac{1}{2}[\alpha(\hat{l}_{sA} - c_s) + \alpha\frac{(1-\hat{\alpha})}{\hat{\alpha}}(l_m - p_{mA}) - \alpha\frac{k_A + c_L}{\hat{\alpha}} + (1-\alpha)(p_{mA} - c_m) + k_A]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} \\
&+\frac{1}{2}[(1-\alpha)\theta(p_{mA} - c_m + k_A)]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} \\
&+\frac{1}{2}[\alpha(\hat{l}_{sB} - c_s) + \alpha\frac{(1-\hat{\alpha}_B)}{\hat{\alpha}_B}(l_m - p_{mB}) - \alpha\frac{k_B + c_L}{\hat{\alpha}_B} + (1-\alpha)(1-\theta)(p_{mB} - c_m) + (\alpha + (1-\alpha)(
\end{aligned}
$$

The doctor can separately set prices for different types of problem.

For type A:

$$
\begin{aligned}
\pi^A &= \frac{1}{2}[\alpha(\hat{l}_{sA} - c_s) + \alpha\frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A}(l_m - p_{mA}) - \alpha\frac{k_A + c_L}{\hat{\alpha}_A} + (1-\alpha)(p_{mA} - c_m) + k_A]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} \\
&+\frac{1}{2}[(1-\alpha)\theta(p_{mA} - c_m + k_A)]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}
\end{aligned}
$$

$$
\begin{aligned}
\hat{l}_{sA} &= \frac{1}{2}[\alpha(\hat{l}_{sA} - c_s) + \alpha\frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A}(l_m - p_{mA}) - \alpha\frac{k_A + c_L}{\hat{\alpha}_A} + (1-\alpha)(p_{mA} - c_m) + k_A]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} \\
&+\frac{1}{2}[(1-\alpha)\theta(p_{mA} - c_m + k_A)]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} + \lambda[\hat{l}_{sA} + \frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A}(l_m - p_{mA}) - \frac{k_A + c_L}{\hat{\alpha}_A} - c_s]
\end{aligned}
$$

$$
\frac{\partial \hat{l}_{sA}}{\partial k_A} = \frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[1 - \frac{\alpha}{\hat{\alpha}_A} + (1-\alpha)\theta] - \frac{\lambda}{\hat{\alpha}_A} \leq 0 \qquad (= 0 \text{ if } k_A > 0)
$$

$$
\begin{aligned}
\lambda[\hat{l}_{sA} + \frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A}(l_m - p_{mA}) - \frac{k_A + c_L}{\hat{\alpha}_A} - c_s] &= 0 \\
\lambda &\geq 0 \\
\hat{l}_{sA} + \frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A}(l_m - p_{mA}) - \frac{k_A + c_L}{\hat{\alpha}_A} - c_s &\geq 0
\end{aligned}
$$

$$\frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[1 - \frac{\alpha}{\hat{\alpha}_A} + (1-\alpha)\theta] - \frac{\lambda}{\hat{\alpha}_A} \leq 0$$

$$\Longrightarrow \frac{\lambda}{\hat{\alpha}_A} \geq \frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[1 - \frac{\alpha}{\hat{\alpha}_A} + (1-\alpha)\theta] > 0$$

$$\Longrightarrow \hat{l}_{sA} + \frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A}(l_m - p_{mA}) - \frac{k_A + c_L}{\hat{\alpha}_A} - c_s = 0$$

$$\Longrightarrow k_A = \hat{\alpha}_A(\hat{l}_{sA} - c_s) + (1-\hat{\alpha}_A)(l_m - p_{mA}) - c_L > 0$$

$$\Longrightarrow \frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[1 - \frac{\alpha}{\hat{\alpha}_A} + (1-\alpha)\theta] - \frac{\lambda}{\hat{\alpha}_A} = 0$$

$$\Longrightarrow \lambda = \frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[\hat{\alpha}_A - \alpha + \hat{\alpha}_A(1-\alpha)\theta]$$

$$\begin{aligned}
\frac{\partial \hat{l}_{sA}}{\partial p_{mA}} &= \frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[-\alpha\frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A} + (1-\alpha) + (1-\alpha)\theta] - \lambda\frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A} \\
&= \frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[-\alpha\frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A} + (1-\alpha) + (1-\alpha)\theta - \frac{(1-\hat{\alpha}_A)}{\hat{\alpha}_A}(\hat{\alpha}_A - \alpha + \hat{\alpha}_A(1-\alpha)\theta)] \\
&= \frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[(1-\alpha) + (1-\alpha)\theta - (1-\hat{\alpha}_A)(1 + (1-\alpha)\theta)] \\
&= \frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[\hat{\alpha}_A - \alpha + \hat{\alpha}_A(1-\alpha)\theta] > 0
\end{aligned}$$

Remark: $\hat{l}_{sA} = \bar{l}_s$ is not optimal. $\hat{l}_{sA} < \bar{l}_s$

Then

$$p_{mA} = l_m$$

$$\Longrightarrow k_A = \hat{\alpha}_A(\hat{l}_{sA} - c_s) - c_L$$

$$\begin{aligned}
\pi^A &= \frac{1}{2}[\alpha(\hat{l}_{sA} - c_s) - \alpha\frac{k_A + c_L}{\hat{\alpha}_A} + (1-\alpha)(l_m - c_m) + k_A]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} \\
&\quad + \frac{1}{2}[(1-\alpha)\theta(l_m - c_m + k_A)]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}
\end{aligned}$$

$$
\begin{aligned}
L^A &= \frac{1}{2}[\alpha(\hat{l}_{sA}-c_s)-\alpha\frac{k_A+c_L}{\hat{\alpha}_A}+(1-\alpha)\,(l_m-c_m)+k_A]\frac{\bar{l}_s-\hat{l}_{sA}}{\bar{l}_s-\underline{l}_s} \\
&+\frac{1}{2}[(1-\alpha)\theta(l_m-c_m+k_A)]\frac{\bar{l}_s-\hat{l}_{sA}}{\bar{l}_s-\underline{l}_s} \\
&+\lambda[\hat{l}_{sA}-\frac{k_A+c_L}{\hat{\alpha}_A}-c_s]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial L^A}{\partial \hat{l}_{sA}} &= \frac{1}{2}\alpha\frac{\bar{l}_s-\hat{l}_{sA}}{\bar{l}_s-\underline{l}_s}-\frac{1}{\bar{l}_s-\underline{l}_s}\frac{1}{2}[\alpha(\hat{l}_{sA}-c_s)-\alpha\frac{k_A+c_L}{\hat{\alpha}_A}+(1-\alpha)\,(l_m-c_m)+k_A] \\
&-\frac{1}{2}[(1-\alpha)\theta(l_m-c_m+k_A)]\frac{1}{\bar{l}_s-\underline{l}_s}+\lambda\leq 0 \quad (=0 \text{ if } \hat{l}_{sA}>\underline{l}_s)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial L^A}{\partial \hat{l}_{sA}} &= \frac{1}{2}\alpha\frac{\bar{l}_s-\hat{l}_{sA}}{\bar{l}_s-\underline{l}_s}-\frac{1}{\bar{l}_s-\underline{l}_s}\frac{1}{2}[\alpha(\hat{l}_{sA}-c_s)-\alpha\frac{k_A+c_L}{\hat{\alpha}_A}+(1-\alpha)\,(l_m-c_m)+k_A] \\
&-\frac{1}{2}[(1-\alpha)\theta(l_m-c_m+k_A)]\frac{1}{\bar{l}_s-\underline{l}_s}+\frac{1}{2}\frac{\bar{l}_s-\hat{l}_{sA}}{\bar{l}_s-\underline{l}_s}[\hat{\alpha}_A-\alpha+\hat{\alpha}_A(1-\alpha)\theta] \\
&= \frac{1}{2}\frac{1}{\bar{l}_s-\underline{l}_s}[\alpha(\bar{l}_s-\hat{l}_{sA})-\alpha(\hat{l}_{sA}-c_s)+\alpha\frac{k_A+c_L}{\hat{\alpha}_A} \\
&-(1-\alpha)\,(l_m-c_m)-k_A-(1-\alpha)\theta(l_m-c_m+k_A) \\
&+[\hat{\alpha}_A-\alpha+\hat{\alpha}_A(1-\alpha)\theta](\bar{l}_s-\hat{l}_{sA})] \\
&= \frac{1}{2}\frac{1}{\bar{l}_s-\underline{l}_s}[-(1-\alpha)\,(l_m-c_m)-\hat{\alpha}_A(\hat{l}_{sA}-c_s)+c_L \\
&-(1-\alpha)\theta(l_m-c_m+\hat{\alpha}_A(\hat{l}_{sA}-c_s)-c_L) \\
&+[\hat{\alpha}_A+\hat{\alpha}_A(1-\alpha)\theta](\bar{l}_s-\hat{l}_{sA})] \\
&= \frac{1}{2}\frac{1}{\bar{l}_s-\underline{l}_s}\{[\hat{\alpha}_A+\hat{\alpha}_A(1-\alpha)\theta](\bar{l}_s-2l^a_{NK\_A}+c_s) \\
&-(1-\alpha)\,(1+\theta)(l_m-c_m)+[1+(1-\alpha)\theta]c_L\} \\
&= \frac{1}{2}\frac{1}{\bar{l}_s-\underline{l}_s}\{\hat{\alpha}_A[1+(1-\alpha)\theta](\bar{l}_s-2l^a_{NK\_A}+c_s+\frac{c_L}{\hat{\alpha}_A}) \\
&-(1-\alpha)(1+\theta)(l_m-c_m)\}
\end{aligned}
$$

If $\hat{l}_{sA}>\underline{l}_s \Longrightarrow \frac{\partial L^A}{\partial \hat{l}_{sA}}=0$

$$
\hat{\alpha}_A[1+(1-\alpha)\theta](\bar{l}_s-2l^a_{NK\_A}+c_s+\frac{c_L}{\hat{\alpha}_A})-(1-\alpha)(1+\theta)(l_m-c_m)=0
$$

$$\begin{aligned}
\hat{l}_{sA} &= \frac{\hat{\alpha}_A[1+(1-\alpha)\theta](\bar{l}_s+c_s+\frac{c_L}{\hat{\alpha}_A})-(1-\alpha)(1+\theta)(l_m-c_m)}{2\hat{\alpha}[1+(1-\alpha)\theta]} \\
&= \frac{\hat{\alpha}_A(\bar{l}_s+c_s)-\frac{(1-\alpha)(1+\theta)}{1+(1-\alpha)\theta}(l_m-c_m)+c_L}{2\hat{\alpha}} \\
&= \bar{l}_s-\frac{\hat{\alpha}_A(\bar{l}_s-c_s)+\frac{(1-\alpha)(1+\theta)}{1+(1-\alpha)\theta}(l_m-c_m)-c_L}{2\hat{\alpha}_A}
\end{aligned}$$

then we need $\bar{l}_s-\frac{\hat{\alpha}_A(\bar{l}_s-c_s)+\frac{(1-\alpha)(1+\theta)}{1+(1-\alpha)\theta}(l_m-c_m)-c_L}{2\hat{\alpha}}>\underline{l}_s$

In this case, $k_A=\hat{\alpha}_A(\hat{l}_{sA}-c_s)-c_L=\frac{\hat{\alpha}_A(\bar{l}_s-c_s)-\frac{(1-\alpha)(1+\theta)}{1+(1-\alpha)\theta}(l_m-c_m)-c_L}{2}$, $p_{sA}=c_s$.

$$\begin{aligned}
\hat{l}^a_{sA} &= \frac{\hat{\alpha}_A(\hat{l}_{sA}-c_s)}{\hat{\alpha}^a_A}+c_s \\
&= \frac{\hat{\alpha}_A(\bar{l}_s-c_s)-\frac{(1-\alpha)(1+\theta)}{1+(1-\alpha)\theta}(l_m-c_m)+c_L}{2\hat{\alpha}^a_A}+c_s \\
\hat{l}^u_{sA} &= \frac{\hat{\alpha}_A(\hat{l}_{sA}-c_s)}{\tilde{\alpha}}+c_s \\
&= \frac{\hat{\alpha}_A(\bar{l}_s-c_s)-\frac{(1-\alpha)(1+\theta)}{1+(1-\alpha)\theta}(l_m-c_m)+c_L}{2\tilde{\alpha}}+c_s
\end{aligned}$$

When cost of lab test is small enough, $\hat{l}_{sA}$ increasing in $\hat{\alpha}_A$ thus decreasing in ω.

Remark: check constraint (**??**):

$$\begin{aligned}
\frac{\bar{l}_s-\hat{l}_{sA}}{\bar{l}_s-\underline{l}_s}(k_A+l_m-c_m) &\geq l_m-c_m \\
\hat{\alpha}_A(\bar{l}_s-\hat{l}_{sA})(k_A+l_m-c_m) &\geq \hat{\alpha}_A(l_m-c_m)(\bar{l}_s-\underline{l}_s)
\end{aligned}$$

since $\hat{\alpha}_A(\bar{l}_s-\hat{l}_{sA})<\hat{\alpha}_A(\bar{l}_s-\underline{l}_s)$

$k_A+l_m-c_m\geq l_m-c_m$

$\hat{\alpha}_A(\bar{l}_s-\hat{l}_{sA})+k_A+l_m-c_m=\hat{\alpha}_A(\bar{l}_s-c_s)+l_m-c_m-c_L>(l_m-c_m)+\hat{\alpha}_A(\bar{l}_s-\underline{l}_s)$

then we have (**??**) holds.

When $\hat{l}_{sA}=\underline{l}_s$,(**??**) automatically holds.

Now we solve pricing problem for type B.

$$\pi_B = \frac{1}{2}[\alpha(\hat{l}_{sB} - c_s) - \alpha\frac{c_L}{\tilde{\alpha}} + (1-\alpha)\,(1-\theta)(l_m - c_m)]\frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s}$$

Then price of minor treatment for problem B will not affect the profit of doctor.

$$\begin{aligned}\frac{\partial \pi_B}{\partial \hat{l}_{sB}} &= \frac{1}{2}\alpha\frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s} - \frac{1}{2}[\alpha(\hat{l}_{sB} - c_s) + (1-\alpha)\,(1-\theta)(l_m - c_m) - \alpha\frac{c_L}{\tilde{\alpha}}]\frac{1}{\bar{l}_s - \underline{l}_s} \\ &= \frac{1}{2}\frac{1}{\bar{l}_s - \underline{l}_s}[\alpha\bar{l}_s - 2\alpha\hat{l}_{sB} + \alpha c_s - (1-\alpha)\,(1-\theta)(l_m - c_m) + \alpha\frac{c_L}{\tilde{\alpha}}]\end{aligned}$$

$$\frac{\partial \pi^u_{NK\_B}}{\partial \hat{l}_{sB}} = 0 \Longrightarrow \alpha\bar{l}_s - 2\alpha\hat{l}_{sB} + \alpha c_s - (1-\alpha)\,(1-\theta)(l_m - c_m) + \alpha\frac{c_L}{\tilde{\alpha}} = 0$$

$$\begin{aligned}\hat{l}_{sB} &= \frac{\alpha\bar{l}_s + \alpha c_s - (1-\alpha)\,(1-\theta)(l_m - c_m) + \alpha\frac{c_L}{\tilde{\alpha}}c_L}{2\alpha} \\ &= \bar{l}_s - \frac{\alpha(\bar{l}_s - c_s) + (1-\alpha)\,(1-\theta)(l_m - c_m) - \alpha\frac{c_L}{\tilde{\alpha}}c_L}{2\alpha} \\ &= \bar{l}_s - \frac{\tilde{\alpha}(\bar{l}_s - c_s) + (1-\tilde{\alpha})(l_m - c_m) - c_L}{2\tilde{\alpha}}\end{aligned}$$

Let $p_{mB} = l_m$, $k_B = 0$ and $p_{sB} = \hat{l}_{sB} - \frac{c_L}{\tilde{\alpha}}$. Remark, given this sepcification, (17) holds

$$\begin{aligned}\pi &= \frac{1}{2}[(1-\alpha)\,(l_m - c_m) + \hat{\alpha}_A(\hat{l}_{sA} - c_s) - c_L + (1-\alpha)\theta(l_{mA} - c_m + \hat{\alpha}_A(\hat{l}_{sA} - c_s) - c_L)]\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s} \\ &\quad + \frac{1}{2}[\alpha(\hat{l}_{sB} - \frac{c_L}{\tilde{\alpha}} - c_s) + (1-\alpha)\,(1-\theta)(l_m - c_m)]\frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s} \\ &= \frac{1}{2}\frac{\bar{l}_s - \hat{l}_{sA}}{\bar{l}_s - \underline{l}_s}[\hat{\alpha}_A((1-\alpha)\theta + 1)(\hat{l}_{sA} - c_s) + (1-\alpha)\,(1+\theta)(l_m - c_m) - ((1-\alpha)\theta + 1)c_L] \\ &\quad + \frac{1}{2}[\alpha(\hat{l}_{sB} - c_s) + (1-\alpha)\,(1-\theta)(l_m - c_m) - \frac{\alpha c_L}{\tilde{\alpha}}]\frac{\bar{l}_s - \hat{l}_{sB}}{\bar{l}_s - \underline{l}_s} \\ &= \frac{1}{2}\hat{\alpha}_A[(1-\alpha)\theta + 1]\frac{(\bar{l}_s - \hat{l}_{sA})^2}{\bar{l}_s - \underline{l}_s} + \frac{1}{2}\alpha\frac{(\bar{l}_s - \hat{l}_{sB})^2}{\bar{l}_s - \underline{l}_s}\end{aligned}$$

$$\begin{aligned}\frac{\partial\pi}{\partial\hat{\alpha}_A} &= \frac{1}{2}[(1-\alpha)\theta+1]\frac{(\bar{l}_s-\hat{l}_{sA})^2}{\bar{l}_s-\underline{l}_s}-\frac{1}{2}\frac{2(\bar{l}_s-\hat{l}_{sA})}{\bar{l}_s-\underline{l}_s}\frac{(1-\alpha)(1+\theta)(l_m-c_m)-[1+(1-\alpha)\theta]c_L}{2\hat{\alpha}} \\ &= \frac{(\bar{l}_s-\hat{l}_{sA})}{\bar{l}_s-\underline{l}_s}\{\frac{1}{2}[(1-\alpha)\theta+1](\bar{l}_s-\hat{l}_{sA})-\frac{(1-\alpha)(1+\theta)(l_m-c_m)-[1+(1-\alpha)\theta]c_L}{2\hat{\alpha}_A}\} \\ &= \frac{(\bar{l}_s-\hat{l}_{sA})}{\bar{l}_s-\underline{l}_s}\frac{[(1-\alpha)\theta+1](\bar{l}_s-\hat{l}_{sA})-(1-\alpha)(1+\theta)(l_m-c_m)+[1+(1-\alpha)\theta]c_L}{2\hat{\alpha}_A}>0\end{aligned}$$

Discussion about corner solution of prices of lab test A:

$$p_{LA}+\hat{\alpha}_A^a p_{sA}+(1-\hat{\alpha}_A^a)p_{mA}\leq\hat{\alpha}_A^a l_s+(1-\hat{\alpha}_A^a)l_m$$

$$p_{LA}+\hat{\alpha}_A^u p_{sA}+(1-\hat{\alpha}_A^u)p_{mA}\leq\hat{\alpha}_A^u l_s+(1-\hat{\alpha}_A^u)l_m$$

$$\hat{l}_{sA}=\omega\hat{l}_{sA}^a+(1-\omega)\hat{l}_{sA}^u$$

$$\pi_A^a=\frac{1}{2}[\alpha(p_{sA}-c_s)+(1-\alpha)(1+\theta)(p_{mA}-c_m)+(1+(1-\alpha)\theta)k_A]\frac{\bar{l}_s-[\omega\hat{l}_{sA}^a+(1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s-\underline{l}_s}$$

$$\begin{aligned}L_A^a &= \frac{1}{2}[\alpha(p_{sA}-c_s)+(1-\alpha)(1+\theta)(p_{mA}-c_m)+(1+(1-\alpha)\theta)k_A]\frac{\bar{l}_s-[\omega\hat{l}_{sA}^a+(1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s-\underline{l}_s} \\ &\quad+\lambda_1[\hat{\alpha}_A^a\hat{l}_{sA}^a+(1-\hat{\alpha}_A^a)l_m-k_A-c_L-\hat{\alpha}_A^a p_{sA}-(1-\hat{\alpha}_A^u)p_{mA}] \\ &\quad+\lambda_2[\hat{\alpha}_A^u\hat{l}_{sA}^u+(1-\hat{\alpha}_A^u)l_m-k_A-c_L-\hat{\alpha}_A^u p_{sA}-(1-\hat{\alpha}_A^u)p_{mA}]\end{aligned}$$

$$\frac{\partial L_A^a}{\partial p_{mA}}=\frac{1}{2}(1-\alpha)(1+\theta)\frac{\bar{l}_s-[\omega\hat{l}_{sA}^a+(1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s-\underline{l}_s}-\lambda_1(1-\hat{\alpha}_A^a)-\lambda_2(1-\hat{\alpha}_A^u)$$

$$\frac{\partial L_A^a}{\partial p_{sA}}=\frac{1}{2}\alpha\frac{\bar{l}_s-[\omega\hat{l}_{sA}^a+(1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s-\underline{l}_s}-\lambda_1\hat{\alpha}_A^a-\lambda_2\hat{\alpha}_A^u$$

$$\frac{\partial L_A^a}{\partial k_A}=\frac{1}{2}(1+(1-\alpha)\theta)\frac{\bar{l}_s-[\omega\hat{l}_{sA}^a+(1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s-\underline{l}_s}-\lambda_1-\lambda_2$$

$$\frac{\partial L_A^a}{\partial \hat{l}_{sA}^a} = \frac{1}{2}[\alpha(p_{sA} - c_s) + (1-\alpha)(1+\theta)(p_{mA} - c_m) + (1 + (1-\alpha)\theta)k_A]\frac{-\omega}{\bar{l}_s - \underline{l}_s} + \lambda_1 \hat{\alpha}_A^a$$

$$\frac{\partial L_A^a}{\partial \hat{l}_{sA}^u} = \frac{1}{2}[\alpha(p_{sA} - c_s) + (1-\alpha)(1+\theta)(p_{mA} - c_m) + (1 + (1-\alpha)\theta)k_A]\frac{-(1-\omega)}{\bar{l}_s - \underline{l}_s} + \lambda_2 \hat{\alpha}_A^u$$

If $\lambda_1 > 0$ and $\lambda_2 = 0$, then $\frac{\partial L_A^a}{\partial \hat{l}_{sA}^u} < 0$ and $\hat{l}_{sA}^u = \underline{l}_s$

$$\lambda_1 = \frac{1}{2}(1 + (1-\alpha)\theta)\frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + (1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s - \underline{l}_s}$$

$$\begin{aligned}
\frac{\partial L_A^a}{\partial p_{mA}} &= \frac{1}{2}(1-\alpha)(1+\theta)\frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + (1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s - \underline{l}_s} \\
&\quad -\frac{1}{2}(1 + (1-\alpha)\theta)\frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + (1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s - \underline{l}_s}(1 - \hat{\alpha}_A^a) \\
&= \frac{1}{2}\frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + (1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s - \underline{l}_s}[(1-\alpha)(1+\theta) - (1 + (1-\alpha)\theta)(1 - \hat{\alpha}_A^a)] \\
&= \frac{1}{2}\frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + (1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s - \underline{l}_s}[(1-\alpha)(1+\theta) - (1 + (1-\alpha)\theta)\frac{(1-\alpha)(1-\theta)}{1 + (1-\alpha)\theta}] = 0
\end{aligned}$$

$$\begin{aligned}
\frac{\partial L_A^a}{\partial p_{sA}} &= \frac{1}{2}\alpha\frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + (1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s - \underline{l}_s} - \frac{1}{2}(1 + (1-\alpha)\theta)\frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + (1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s - \underline{l}_s}\hat{\alpha}_A^a \\
&= \frac{1}{2}\frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + (1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s - \underline{l}_s}[\alpha - (1 + (1-\alpha)\theta)\hat{\alpha}_A^a] = 0
\end{aligned}$$

$$\lambda_1 = \frac{1}{2}(1 + (1-\alpha)\theta)\frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + (1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s - \underline{l}_s}$$

$$\hat{\alpha}_A^a \hat{l}_{sA}^a + (1 - \hat{\alpha}_A^a)l_m - k_A - c_L - \hat{\alpha}_A^a p_{sA} - (1 - \hat{\alpha}_A^a)p_{mA} = 0$$

$$\begin{aligned}
k_A &= \hat{\alpha}_A^a \hat{l}_{sA}^a + (1 - \hat{\alpha}_A^a)l_m - c_L - \hat{\alpha}_A^a p_{sA} - (1 - \hat{\alpha}_A^a)p_{mA} \\
(1 + (1-\alpha)\theta)k_A &= \alpha \hat{l}_{sA}^a + (1-\alpha)(1+\theta)l_m - (1 + (1-\alpha)\theta)c_L - \alpha p_{sA} - (1-\alpha)(1+\theta)p_{mA}
\end{aligned}$$

$$\frac{\partial L_A^a}{\partial \hat{l}_{sA}^a} = \frac{1}{2}[\alpha(p_{sA} - c_s) + (1-\alpha)\,(1+\theta)(p_{mA} - c_m) + (1 + (1-\alpha)\theta)k_A]\frac{-\omega}{\bar{l}_s - \underline{l}_s} + \lambda_1 \hat{\alpha}_A^a = 0$$

$$\frac{1}{2}[\alpha(p_{sA} - c_s) + (1-\alpha)\,(1+\theta)(p_{mA} - c_m) + (1 + (1-\alpha)\theta)k_A]\frac{\omega}{\bar{l}_s - \underline{l}_s} = \lambda_1 \hat{\alpha}_A^a$$

$$\begin{aligned}
\frac{1}{2}[\alpha(p_{sA} - c_s) + (1-\alpha)\,(1+\theta)(p_{mA} - c_m) + (1 + (1-\alpha)\theta)k_A]\frac{\omega}{\bar{l}_s - \underline{l}_s} &= \frac{1}{2}(1 + (1-\alpha)\theta)\hat{\alpha}_A^a \frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + \cdots}{\bar{l}_s - \cdots} \\
\frac{1}{2}[\alpha(p_{sA} - c_s) + (1-\alpha)\,(1+\theta)(p_{mA} - c_m) + (1 + (1-\alpha)\theta)k_A]\frac{\omega}{\bar{l}_s - \underline{l}_s} &= \frac{1}{2}\alpha \frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + (1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s - \underline{l}_s} \\
\frac{1}{2}[\alpha(\hat{l}_{sA}^a - c_s) + (1-\alpha)\,(1+\theta)(l_m - c_m) - (1 + (1-\alpha)\theta)c_L]\frac{\omega}{\bar{l}_s - \underline{l}_s} &= \frac{1}{2}\alpha \frac{\bar{l}_s - [\omega \hat{l}_{sA}^a + (1-\omega)\hat{l}_{sA}^u]}{\bar{l}_s - \underline{l}_s} \\
\omega[\alpha(\hat{l}_{sA}^a - c_s) + (1-\alpha)\,(1+\theta)(l_m - c_m) - (1 + (1-\alpha)\theta)c_L] &= \alpha \bar{l}_s - \alpha[\omega \hat{l}_{sA}^a + (1-\omega)\underline{l}_s]
\end{aligned}$$

$$2\omega\alpha\hat{l}_{sA}^a = \alpha\bar{l}_s - \alpha(1-\omega)\underline{l}_s - \omega[-\alpha c_s + (1-\alpha)\,(1+\theta)(l_m - c_m) - (1 + (1-\alpha)\theta)c_L]$$

$$\hat{l}_{sA}^a = \frac{\hat{\alpha}_A^a(\frac{\bar{l}_s - (1-\omega)\underline{l}_s}{\omega} + c_s) - \frac{(1-\alpha)(1+\theta)}{1+(1-\alpha)\theta}(l_m - c_m) + c_L}{2\hat{\alpha}_A^a}$$