The Unintended Consequences of Clients' Active Participation in Expert Services: Theory and Experiment*

Nanyin Yang†

April 20, 2024

Click here for the latest version.

Abstract

In expert services, such as medical services, clients often lack specialized skills to diagnose their problems. However, it is common for them to participate in decision making through self-diagnosis, potentially leading to noncompliance against experts' advice. This study investigates how clients' active participation impacts experts' investment of effort in the precision of diagnosis. In a theory-driven lab experiment, passive and active clients, who vary in whether they have an option to go against experts, consult experts for diagnosing unknown problems, and experts exert effort to improve the diagnostic precision. The results show that giving clients an option to go against experts' advice reduces experts' investment in improving the precision of diagnosis. This effect is particularly large among experts who prioritize clients' well-being. Additionally, providing clients with more information about experts, such as through a rating system, mitigates the negative effects of client participation. This study sheds light on the potential adverse effect of client involvement in expert services.

Keywords: Expert services, Client Noncompliance, Lab experiments
JEL Codes: C91, I11

*I would like to especially thank my advisor Catherine Eckel for her guidance and support. I thank Loukas Balafoutas, Jeannette Brosig-Koch, Ben Bushong, Li Gan, Huiyi Guo, Rudolf Kerschbamer, Silvana Krasteva, Joanna Lahey, Joanathan Meer, Marco Palma, Danila Serra, and numerous seminar participants for helpful comments. This study is approved by the Institutional Review Boards at TAMU (IRB2022-0765M). The experiment is pre-registered on AsPredicted (No. 105234). This study is funded by the National Science Foundation under Grant No. 2215032.

†Department of Economics, Texas A&M University

1 Introduction

Expert service differs from other services in that clients lack a clear understanding of their needs: they rely on experts' specialized skills to diagnose problems and provide treatments. Though lacking expertise in diagnosis, in real life it is common for clients to actively participate through self-diagnosis, sometimes resulting in noncompliance against their experts. Consider medical service as an example. Although the patient-doctor relationship has been considered using a top-down paternalist model (Emanuel and Emanuel, 1992), nowadays, more and more patients actively involve themselves in medical decision-making, partly driven by the democratization of information through the internet (Tan and Goonawardene, 2017). Using a nationwide survey on internet usage, Fox and Duggan (2013) reported that around 35% of U.S. adult internet users used the internet as a diagnostic tool, with 18% of them reported self-diagnoses conflicting with doctors' diagnoses. Those low-cost access to health information, though not necessarily accurate, empowers patients to be actively involved in medical decision making.

It remains unclear how clients' active participation changes experts' behaviors. In the medical service context, given that doctors commonly hold the goal of solving patients' health problems, patients' active participation may be perceived as potential noncompliance against doctors and discourage doctors from exerting more effort for a more accurate diagnosis. Some empirical evidence provides support for this concern: In a national survey, 59% of doctors reported "patient's request" as the main reason for over-treatment in the U.S. (Lyu et al., 2017). Moreover, physician burnout is positively associated with higher frequency of difficult patient encounters (An et al., 2013), though the causal direction is not clear. Due to the challenges in empirical study, there is lack of a direct causal evidence on how clients' active participation and potential noncompliance affect experts' efforts in diagnosis and treatment giving.

In this study, I discuss the influence of clients' active participation on experts' investment of effort in the diagnosis and treatment giving. In my theory-driven lab experiment, a client consults an expert to diagnose an unknown problem and recommend a treatment. Following Balafoutas et al. (2020), experts decide the investment in improving the diagnostic precision, which determines the probability of obtaining a correct diagnosis. This setup captures in real-life scenarios, experts' diagnosis may be associated with uncertainty (Beresford, 1991). Clients, though lacking expertise in problem diagnosis, have an inaccurate self-diagnosis which may change their decisions on whether to follow their experts' recommendations.

To understand the influence of clients' active participation on experts' investment in diagnostic precision, within each experimental session, I introduce two types of clients, "passive" and "active". The distinction lies in whether the client has an option to overrule experts' recommendations on treatments. Experts within each session vary in their "incentive alignment" with the clients, varying their concerns/motivations to solve clients' problems. Furthermore, I impose two between-subject treatments: under the "Observable" condition, the expert's diagnostic precision is observed by the client; under the "Concealed" condition, the expert's diagnostic precision is not observable, which is closer to real life where experts' effort and diagnostic precision are challenging to gauge.

The experimental results show that the Concealed condition leads to experts' differing attitudes toward active and passive clients. When the diagnostic precision is observable to clients, experts with high incentive alignment invest equally in active and passive clients. However, when the diagnostic precision is not observable, experts invest less in diagnostic precision for active clients than for passive clients. This investment gap increases with experts' incentive alignment with clients. In other words, experts with stronger motivation to solve clients' problems exhibit a greater difference in investment between active and passive clients. As a result, giving clients an option to overrule experts results in lowering their probability of solving problems. In the realm of medical services, this finding implies that patients' active participation in treatment decisions may lead to worse health outcomes. When patients cannot assess the accuracy of doctors' diagnosis, doctors may perceive their participation in decision-making as potential noncompliance. Doctors who prioritize patients' well-being are more likely to be concerned about wasting efforts, leading to reduced diagnostic effort. Consequently, compared with patients who always fully follow doctors, patients given the option to challenge doctors receive less diagnostic effort and face a higher risk of unresolved medical issues.

In addition to the Observable and Concealed conditions, I further explore two institutional changes: Communication and Reputation conditions. Under Observable and Concealed conditions, clients remain unaware of any information about experts to maintain controlled testing environments. However, real-world interactions often involve clients having prior information about experts, aiding their decision-making. Therefore, I introduce Communication (direct conversation with experts) and Reputation (experts' public ratings) conditions, mimicking how clients typically gather information about experts. Results show that both conditions reduce the investment gap for active and passive clients and increase problem-solving probability. Both conditions also increase the probability of solving problems. However, the increase in problem solving does not rely on increased diagnostic precision. Instead, they improve clients' well-being by providing information to help clients make better decisions. The Communication condition is close to the Observable condition – clients form accurate beliefs about experts' diagnostic precision, allowing them to follow highly precise diagnoses and disregard imprecise ones. Under the Reputation condition, clients' ratings of experts effectively inform clients' compliance choices.

This study adds to the expert service literature by studying the adverse effect of clients' active participation on experts' efforts in diagnosis. The experimental results suggest that when experts are motivated to solve clients' problems but are not able disclose their diagnostic precision, clients' active involvement in decision-making discourages experts from investing more effort in diagnosis. This finding is important as inferring the causal effect of clients' involvement on experts' effort from natural data is challenging. Furthermore, the study shows that experts who are more closely aligned with their clients' incentives have different attitudes depending on client compliance. In the medical context, one can interpret this incentive alignment as the healthcare provider's prosocial concerns about patients' health. One can also consider different levels of incentive alignment as a spectrum of expert-service contracts, varying in how closely expert incentives align with client well-being. This study implies that in industries where solving clients' problems directly impacts experts' welfare, experts are more sensitive about clients' trust and compliance.

In addition, by imposing two exploratory institutional changes, this study suggests that giving clients more information on experts helps reduce the adverse effect of client participation. In particular, the Reputation condition offers clients an indirect way to learn about their experts. Clients, even if with no access to experts' diagnostic precision or incentive alignment, can still infer from other clients' ratings about whether an expert's diagnosis is precise enough to be followed. Compared with direct conversations, the online rating system serve as a less costly way to improve clients' well-being, even if it does not entirely eliminate experts' investment gap between active and passive clients.

2 Related Literature

This study is closely related to the expert service literature. Previous studies on clients' active participation in expert services and expert's other-regarding preferences are discussed in this subsection. In addition, under the medical context, this study complements the empirical studies on healthcare provider behaviors and adds to the literature on experimental health economics.

Expert services, where experts know more than clients about clients' needs for services, are first studied by Darby and Karni (1973). In these so-called credence good markets, such as healthcare and car repair services, Dulleck and Kerschbamer (2006) predicts that when clients fully commit to experts' recommended treatments, the information asymmetry can result in market failure (no trade at all). Subsequent research has examined the impact of clients' participation on the market equilibrium, including clients' rejection of expert advice, seeking of second opinion, and personal research on diagnosis. Those studies are mixed on the impact of clients' participation. Mimra et al. (2016) demonstrate through a laboratory experiment that access to costly second opinions significantly reduces experts' overtreatment. As pointed out by Balafoutas and Kerschbamer (2020), improved access to information is beneficial in improving clients' autonomy and reducing experts' dishonesty. Conversely, there are also studies exploring the potential drawbacks of client autonomy. For example, Fong et al. (2014) consider a scenario where clients can reject experts' recommendations. They find that if clients do not commit to experts' recommendations, the market can end up with inefficiency even if the service quality is verifiable. Furthermore, Dulleck and Kerschbamer (2009) combines the assumptions of expert diagnostic uncertainty and consumer non-commitment to study the clients' free-riding on expert diagnoses – clients may obtain a diagnosis from experts and then switch to cheaper discounters for actual treatments. They show that such free-riding discourages experts from exerting higher efforts in diagnosis and can lead to undertreatment equilibria.[1] My study extends previous findings by showing how clients' active participation could discourage experts from providing better services, especially when experts are highly motivated to help clients solve their problems.

In addition, there are two important settings discussed in the expert service literature, which are closely related to this study: diagnostic uncertainty, and expert's other-regarding preferences. In recent years, there have been a few studies discussing the role of diagnostic uncertainty, as exogenously determined by experts' ability or endogenously determined by experts' investment of effort. Hilger (2016) extends the model developed by Dulleck and Kerschbamer (2006), demonstrating that when the diagnostic cost is not observable, mistreatment always exists in equilibrium. Liu et al. (2020) investigates the market equilibrium where experts are heterogeneous in their diagnostic ability. Their theoretical results imply that under certain settings, a higher proportion of high-ability experts in the market can reduce market efficiency. Balafoutas et al. (2020) discuss the role of diagnostic uncertainty in a framework that considers insurance coverage. They found that insurance reduces experts' investment in the diagnosis and lowers the diagnostic precision. The theoretical framework of this study borrows the setup from Balafoutas et al. (2020) by assuming that experts have to exert costly effort to improve the precision of the diagnosis.

Regarding experts' other-regarding preferences, Kerschbamer et al. (2017) designed an

[1] Clients' participation in decision-making can also occur in earlier stages of the service. For example, Schulte and Felgenhauer (2017) study the impact of clients' participation in the pre-selection of investment projects. They find that if a client pre-selects a project for an expert to evaluate, the expert will know that the client favors this project and will be biased towards recommending executing this project. As a result, clients' involvement in the selection of projects increases the risk of investing in a bad project.

experiment that enables a clear discrimination of experts' heterogeneous social preferences. They show that social preferences provide a better explanation of experts' behaviors in the experiment than the standard assumption of selfish preferences. In the experiment by Balafoutas et al. (2020), experts' social preferences also play an important role in improving the diagnostic precision. In my theoretical framework, I use the term "incentive alignment" to generalize the experts' other-regarding preferences into the alignment of motivation on problem solving between experts and clients. I add to the expert service literature by showing that incentive alignment between experts and clients plays a vital role in experts' perspectives of clients' active participation – intuitively, experts who are more concerned about clients' well-being are more sensitive about clients' compliance.

By giving clients some noisy private information, clients in the experiment can actively participate in diagnosis and go against experts' recommendations. Previous literature is divided on the impact of clients' active participation. As pointed out by Balafoutas and Kerschbamer (2020), improved access to information is beneficial in improving clients' autonomy and reducing experts' dishonesty. There are also prior studies exploring the potential downside of client autonomy. For example, Fong et al. (2014) consider a scenario where clients are able to reject experts' recommendations. They find that if clients do not commit to experts' recommendations, the market can end up with inefficiency even if the service quality is verifiable. Furthermore, Dulleck and Kerschbamer (2009) combines the assumptions of expert diagnostic uncertainty and consumer non-commitment to study the clients' free-riding on expert's diagnosis – clients may ask for a diagnosis from experts and then switch to cheaper discounters for actual treatments, which discourages experts from exerting higher efforts in diagnosis. [2] My study extends previous findings by showing how clients' autonomy could play a discouraging role when experts are highly motivated to solve clients' problems: when experts are more incentive-aligned with clients, they are more likely to be discouraged by clients' potential noncompliance. As a result, compared with clients who are fully compliant, clients with opportunities to go against experts receive lower efforts in diagnosis from experts.

The findings of this paper can be applied to healthcare services. There is a small but growing body of literature using laboratory experiments to study healthcare-related topics, for example, the design of incentive schemes (Brosig-Koch et al., 2016, 2017; Hennig-Schmidt et al., 2011). As indicated by Galizzi and Wiesen (2018), there is a growing acceptance of

[2] Clients' participation in decision-making can also occur in earlier stages of the service. For example, Schulte and Felgenhauer (2017) study the impact of clients' participation in the pre-selection of investment projects. They find that if a client pre-selects a project for an expert to evaluate, the expert will know that the client favors this project and will be biased towards recommending executing this project. As a result, clients' involvement in the selection of projects increases the risk of investing in a bad project.

the laboratory approach in health economics, coinciding with the application of behavioral economics to health research. My experimental findings complement previous empirical studies on patient-doctor relationships. Previous empirical studies show that doctors do not treat patients in identical ways. Their medical treatments differ based on patients' education levels (Brekke et al., 2018), socio-economic status (Banuri et al., 2018; Gottschalk et al., 2020), ethnicity (Alsan et al., 2019; Schulman et al., 1999), gender (Schulman et al., 1999), etc. My study adds to those discussions by suggesting that patients' active involvement and potential noncompliance can also change doctors' service quality, which to my best knowledge is not fully explored in the literature.

3 Theoretical Framework: Expert-Client Interaction

3.1 Basic Setup

I consider an economy with one client and one expert. The client is facing an unknown problem, z, and has to visit an expert for diagnosis and treatment to solve the problem. To simplify the discussion, I assume that there are two problems $z \in \{0, 1\}$, and the ex-ante probability of having problem $z = 1$ is 0.5, i.e. these two problems are equally likely to occur. Furthermore, there are two treatments $T \in \{0, 1\}$ that can solve each of these problems: treatment $T = 1$ solves $z = 1$, and treatment $T = 0$ solves $z = 0$.

The client receives a private signal $s^c \in \{0, 1\}$ about the problem, with the precision $Pr(s^c = z) = q$, where $q \in [0.5, 1]$. This private signal is analogous to patients' self-diagnoses in real life, which could come from online searches, and are not necessarily correct. The expert receives a diagnostic signal $s^x \in \{0, 1\}$ that correctly identifies the problem with probability E, i.e. $Pr(s^x = z) = E$. In line with Balafoutas et al. (2020), I assume that the expert can freely choose the level of diagnostic precision $E \in [0.5, 1]$, associated with an effort cost $g(E) = k(E - 0.5)^2$, where k is a strictly positive parameter. After the expert conducts the diagnosis, the diagnostic result is automatically converted to a treatment recommendation $T^x \in \{0, 1\}$, with $T^x = s^x$ (i.e. I assume that the expert is not able to tell lies by providing the treatment that is inconsistent with the problem that they diagnose). This setting is similar to the scenario where the higher the effort a doctor invests in a patient's case, the more precise the diagnostic result will be.[3] After getting a diagnosis, the doctor follows a

[3] When $E = 0.5$, it means that the expert's diagnosis is a 50/50 random draw which does not convey any useful information about the problem. One may be concerned that in real life, doctors usually do not invest so little in the diagnosis because of the fear of medical malpractice accusations. I believe that this simplification is worthwhile, because in the experiment, what I am focusing on is the change in the doctor's diagnostic precision between treatments rather than the absolute level of diagnostic precision. Future research could introduce punishment for malpractice.

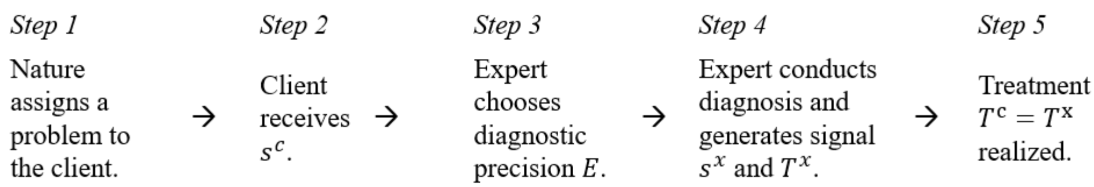


Figure 1: Game Procedure with Passive Client

guideline on giving treatments, without any room to manipulate the interpretation of the diagnostic result or to purposely mislead patients to another treatment.

To simplify the discussion, I assume that the client's goal is to solve the problem. Therefore, the client's utility is only determined by whether the treatment that the client actually receives, namely T^c, is the desired treatment: $U^c = H$ if $T^c = z$, and $= L$ if $T^c \neq z$ (assume: H and L are constants, and $H > L$). Following Balafoutas et al. (2020) and Liu et al. (2020), I assume that the expert cares about the client's health status with an incentive alignment parameter $\gamma \in [0, 1]$. The expert is under a capitation payment system, so he receives a lump sum M for serving one client. Thus, the expert's utility U^x is defined in the following way: $U^x = M + \gamma(U^x - L) - g(E)$. Notice that the term $U^c - L$ equals $H - L$ if the problem is solved and 0 if the problem is not solved. Therefore, the incentive alignment γ determines the extra utility that the expert receives from improving the client's well-being. In other words, this alignment parameter γ captures the expert's concerns about the client's well-being. One could interpret this alignment in various ways, for example, the expert's inherent altruism, the expert's concerns about their own reputation, the expert's fear of malpractice, etc.

I assume two types of clients, *passive* clients, and *active* clients. A passive client always accepts the expert's recommended treatment, mirroring the traditional paternalistic model of patient-doctor relationships, where doctors are the primary decision-makers and patients fully comply with doctors Emanuel and Emanuel (1992). An active client has the option to actively participate in treatment selection, i.e. they can freely go against their experts by choosing the opposite treatment.

In the rest of this section, I will discuss the interaction between the client and the expert in three cases. The first case ("passive client") is a benchmark case, where the client is passive through the whole interaction process, which is analogous to the traditional paternalistic model of patient-doctor relationships discussed by Emanuel and Emanuel (1992). Then I further investigate two cases with active clients. In Case 1 ("active client & observable precision"), the client directly observes the expert's diagnostic precision, and determines whether to follow the expert's recommendation. In Case 2 ("active client & concealed precision"), the client does not observe the expert's diagnostic precision, but is still able to overrule the expert's recommendation. In this case, I assume that the client forms a belief about the expert's diagnostic precision, with the expectation of the belief denoted as $\hat{E}$. By comparing the passive-client case with the two active-client cases, I will be able to investigate how clients' active participation changes experts' behaviors, and how the observability of diagnostic precision plays a role in this relationship.

3.2 Benchmark: Passive Client

As mentioned above, the benchmark case assumes that the client always fully complies with the expert. Figure 1 shows the game procedure. First, the nature assigns an unknown problem to the patient. Next, the client receives a private signal s^c about the problem, and the expert determines the effort level E. After investing the diagnostic effort, a diagnostic result s^x is drawn which is correct with probability E. The diagnostic result s^x will be converted into a treatment T^c for the patient, and both parties' utilities are realized. [4]

The expected utility for the doctor is:

$$E(U^x) = M + \gamma(EH + (1-E)L) - d(E-0.5)^2 \tag{1}$$

Then, the optimal effort level for the doctor will be:

$$E^{passive} = g'^{(-1)}(\gamma(H-L)) \tag{2}$$

where $g'^{(-1)}(\cdot)$ is the inverse function of the first derivative of the cost function $g(\cdot)$.

Proposition 1 *The expert's optimal precision level for a passive client is:* $E^{passive} = 0.5 + \gamma(H-L)/2k$.

Proof. Take the partial derivatives with respect to E, then the solution to the optimal precision is obvious. ■

Proposition 1 indicates that when the expert knows that the client is fully compliant with the expert's recommendation, the expert's investment of efforts in the diagnostic precision increases with the alignment parameter γ. Notably, when $\gamma = 0$, indicating no concern for the client's well-being, the doctor will not invest in improving the diagnostic precision, i.e. $E^{passive} = 0.5$ and $g(E^{passive}) = 0$.

[4] Notice that in this case, the client is passive in the whole process. Thus, whether the client receives a private signal or observes the doctor's investment in diagnosis does not affect the equilibrium. I still include these steps in Figure 1 in order to keep the game procedure consistent with the experimental design, which will be explained in the next section.

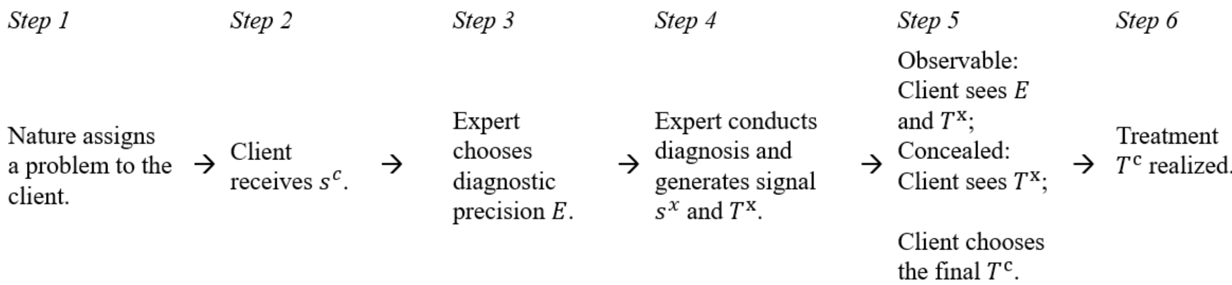


Figure 2: Game Procedure with Active Patient: Observable and Concealed Cases

3.3 Case 1: Active Client & Observable Precision

In this case for discussion, the client is active and involved in choosing treatments. Additionally, the expert's diagnostic precision is observable to the client. Figure 2 describes the timing of the game for this case. Different from the benchmark case, after the expert makes the treatment recommendation in Step 4, in Step 5, the client observes the diagnostic precision E and the recommendation T^x, and then determines the actual treatment T^c to receive. Therefore, the client has the option to go against the expert by choosing the treatment opposite to the recommended one.

Lemma 1 *When $s^c = s^x$, the client always follows the doctor's treatment recommendation regardless of the precision level E.*

Proof. Obvious. ■

Lemma 1 is intuitive: when the client and the expert hold consistent opinions about the problem, there is no reason for the client to overrule the expert's recommendation. The client will consider overruling her expert only when her private signal about the problem conflicts with the expert's diagnosis.

Lemma 2 *When $s^c \neq s^x$, the client follows the expert's treatment recommendation only when $E \geq q$. Otherwise, the patient will revise the treatment into $T^c = s^c$.*

Proof. When $s^c \neq s^x$, the patient's expected utility from following the doctor is: $E^{follow}(U^p) = EH + (1-E)L$. If not following the doctor, the expected utility is: $E^{overrule}(U^p) = qH + (1-q)L$. Since both E and q are between 0.5 and 1, and $H > L$, $E^{follow}(U^p) \geq E^{overrule}(U^p)$ only when $E \geq q$. ■

Lemma 2 describes the client's best strategy in response to the observed diagnostic precision E. It claims that when the expert's diagnostic result conflicts with the client's private signal, the client's compliance depends on the expert's diagnostic precision. If the expert's diagnosis is less precise than the client's privation signal, then the client will choose not to comply.

Proposition 2 *(1) For $\gamma < 4k(q-0.5)/(H-L)$, the expert's equilibrium strategy is $E = 0.5$, and the client will only follow her own private signal on this equilibrium path. (2) For $\gamma \geq 4k(q-0.5)/(H-L)$, the expert's equilibrium strategy is $E^{act\&obs} = 0.5 + \gamma(H-L)/2k$, and the client will always follow the expert on this equilibrium path.*

Proof. To have the expert choose a diagnostic precision level greater than 0.5, two conditions must be satisfied: (i) The expert's expected utility is higher with this precision level than with $E = 0.5$; (2) The client is compliant with the expert. Notice that if the client is fully compliant, then it goes back to the passive client case, so the candidate optimal choice for the expert under this case is $E^{act\&obs} = 0.5 + \gamma(H-L)/2k$. To satisfy (i), I need: $E[U^x(E^{act\&obs})] \geq E[U^x(0.5)]$, which can be simplified to $\gamma \geq 4k(q-0.5)/(H-L)$. Now still need to show: $E^{act\&obs} > q$. Since $\gamma \geq 4k(q-0.5)/(H-L)$, there will be: $E^{act\&obs} > 0.5 + [4k(q-0.5)/(H-L)](H-L)/2k = 2q - 0.5$. Because $q \geq 0.5$, $E^{act\&obs} > q$ always hold with $\gamma \geq 4k(q-0.5)/(H-L)$. ■

According to Proposition 2, the equilibrium depends on γ, i.e. the expert's concern about the client's well-being. If the expert's concern is lower than the threshold (i.e., $\gamma < 4k(q-0.5)/(H-L)$), the market reaches a "low-precision-low-compliance" equilibrium. If the expert has high enough concerns about the client, he will exert sufficient effort to increase the precision. The client observes the precision and fully complies. In this case, the market reaches a so-called "high-precision-high-compliance" equilibrium. This equilibrium outcome mirrors the benchmark case where the client is passive.

3.4 Case 2: Active Client & Concealed Precision

In this section, I discuss the case when the client does not observe the expert's diagnostic precision, but still decides whether to comply with the recommendation. This case is closer to real life, where a patient cannot assess how much effort a doctor invests in the diagnosis, nor the precision of the diagnosis; however, after receiving the expert's diagnosis, the client still must decide whether to follow the recommendation. The procedure is identical to Case 1 with observable precision as shown in Figure 2, except that in Step 5, the client sees only the recommendation T, and chooses whether to comply.

Notice that in this case, Lemma 1 still holds: when the client's private signal is consistent with the expert's diagnosis, then the client's optimal choice is to follow the expert's treatment recommendation. However, in this case, since the client cannot observe the expert's diagnostic precision, when her private information conflicts with the expert's diagnosis, her compliance will depend on the expectation of her belief, $\hat{E}$, about the precision level.

Lemma 3 *When $s^d \neq s^p$, the client follows the expert's treatment recommendation only when $\hat{E} \geq q$. Otherwise, the client will revise the treatment into $T = s^p$.*

Proof. Obviously similar to the proof of Lemma 2. ■

Lemma 3 is the client's best strategy in the case when the diagnostic precision is not observable and the private signal is not consistent with the expert's diagnostic result. It predicts that the client's compliance depends on her belief $\hat{E}$: if the client believes that her private information is more precise than the doctor's diagnosis, she would rather follow her own private signal for choosing the treatment. This lemma is analogous to the real-life scenario where patients do not follow their doctors' suggestions when they believe that their doctors are not investing enough efforts in the diagnosis (e.g. doctors do not spend enough time examining all symptoms or analyzing the test results), and thus believe that their doctors' diagnoses are incorrect. Note that in this case, the expert's optimal choice will not be solely affected by the patient's private signal accuracy q, but also the relationship between q and $\hat{E}$:

Proposition 3 *(1) For $\gamma < 4k(q-0.5)/(H-L)$, the expert's equilibrium strategy is $E = 0.5$, and on this equilibrium path, the client only follows her private signal and chooses $T^c = s^c$. (2) For $\gamma \geq 4k(q-0.5)/(H-L)$, there will be two equilibria: (i) the client believes that $\hat{E} < q$, and on this equilibrium path, the expert chooses $E = 0.5$ while the client only follows her private signal and chooses $T^c = s^c$; (ii) the client believes that $\hat{E} \geq q$, and on this equilibrium path, the expert chooses $E^{act\&conceal} = 0.5 + \gamma(H-L)/2k$ while the client always follows the expert, i.e., chooses $T^c = T^x$.*

Proof. For $\gamma < 4k(q-0.5)/(H-L)$, there will be $E[U^x(0.5+\gamma(H-L)/2k)] < E[U^x(0.5)]$, so the expert will always choose $E = 0.5$, while the client's equilibrium strategy will be only following her private signal. For $\gamma \geq 4k(q-0.5)/(H-L)$, if the client believes that $\hat{E} < q$, then they will not follow the expert, and the doctor's equilibrium strategy is $E = 0.5$. Both parties have no motivations to deviate. If the client believes that $\hat{E} \geq q$, then the expert's equilibrium strategy is $E^{act\&conceal} = 0.5 + \gamma(H-L)/2k$ because the client is now "passive", similar to the case in Proposition 1. ■

Proposition 3 is different from Proposition 2 in that when the diagnostic precision is concealed to the client, the model predicts two equilibria for an expert with alignment parameter $\gamma \geq 4k(q-0.5)/(H-L)$: In one, the expert invests to increase the diagnostic precision, and the client follows the expert; in the other, the expert never invests in improving diagnostic precision, and the client never follows the expert. Notice that the first equilibrium returns identical outcomes as the equilibrium predicted in Proposition 2 (the Active Client &

Observable Precision case) as well as in Proposition 1 (the Passive Client case). The latter is an additional equilibrium that does not exist in either the Active Client & Observable Precision case or the Passive Client case. This additional equilibrium suggests a possible scenario in which even if an expert is highly incentive-aligned with a client, they still end up with a low-precision-low-compliance outcome. Intuitively, this is because if experts are highly motivated to solve problems, they may perceive clients' active participation as possibly ignoring their advice, making their efforts a waste. When the market converges to this equilibrium, compared to Case 1 and the benchmark case, clients in Case 2 receive fewer efforts from experts and are less likely to solve problems.

In summary, in this section, I discuss the patient-doctor relationship using a general expert-client framing. I demonstrate that when the client is not able to assess the precision of the expert's diagnosis (or in other words, the expert's effort in the diagnosis), there will be two equilibria coexisting: one returns a high-precision-high-compliance outcome, another returns a low-precision-low-compliance outcome. Compared with the case when clients are fully compliant (i.e. passive client case), if the latter equilibrium occurs, giving clients an opportunity to overrule experts will make both parties worse off.

However, this theoretical framework does not provide criteria to discuss the likelihood of the second equilibrium to occur. Therefore, I conduct an experiment to further investigate the existence of this equilibrium. In the next section, I will introduce the experimental design motivated by this model, with certain parameter setups. Based on the parameter setups, I will present the testable predictions.

4 Experimental Design

In this section, I first introduce the model parameters for the experiment. The choice of parameters is motivated by the research interest in verifying the impact of client's active participation in treatment selection on the expert's performance. Then I present the treatment conditions and further details of the experimental procedure. In the last part of this section, I discuss predictions based on the experimental parameters. Apart from verifying the theoretical prediction, another goal of this experiment is to investigate the effectiveness of providing information about the experts for clients in increasing the expert's diagnostic precision and improving the client's well-being, which will be presented in the treatment subsection.

Table 1: The Doctor's Cost Table of Diagnostic Precision

Precision	50%	60%	70%	80%	90%	100%
Cost	0	1	4	9	16	25

Note: This is the cost table that subject experts observed when making the precision choice in each round.

4.1 Parameters

In this experiment, each participant is randomly assigned to the role of client or expert, and is randomly matched into pairs. There are four experimental treatment conditions (details will be explained later), and for all treatment conditions, it is common knowledge that each client receives one of two problems (problem A or problem B) with equal probability. The client does not know the problem but receives a signal of either A or B, with the probability of 0.6 that this signal is consistent with her true problem ($q = 0.6$). The diagnostic result and treatment will also be either A or B. If the client receives the correct treatment, she will get 120 tokens ($H = 120$), and if not, she will get 20 tokens ($L = 20$)[5].

The expert gets a lump sum of 80 tokens for each interaction with the client ($M = 80$), with the purpose of balancing the expected payoff between experts and clients. The expert can freely choose the diagnostic precision $E \in \{0.50, 0.60, 0.70, 0.80, 0.90, 1.00\}$. The cost of the diagnosis is $g(E) = 100(E - 0.5)^2$. To simplify the task, a cost table will be presented to the expert as shown in Table 1, where the expert just needs to choose among the six available precision levels. In addition, I randomly assigned four levels of $\gamma \in \{0, 0.2, 0.6, 1\}$ to each expert and this value will be fixed for the whole session for each expert.[6] Therefore, experts with $\gamma = 1$ earn 100 tokens from solving a problem, making them highly motivated to solve the problem. In constrast, experts with $\gamma = 0$ earn 0 tokens from solving a problem, indicating no alignment with clients' incentives in problem-solving.

4.2 Treatment Conditions

There are four treatment conditions in this experiment: *Baseline*, *Concealed condition*, *Communication condition*, and *Reputation condition*, and each persists for 20 rounds.

First of all, to examine the role of clients' active participation in decision-making, I vary the opportunity for clients to choose treatments on their own. In each session, clients will be randomly assigned to be "active" with a probability of 70% and be "passive" with a

[5]The exchange rate is 10 tokens = 1 dollar.

[6]In the experiment, to simplify the meaning of γ, I explain it to subjects by the term "contract". I frame each of the γ values as "0-Contract", "20-Contract", "60-Contract", and "100-Contract", which stand for the extra tokens that an expert will earn from solving the client's problem.

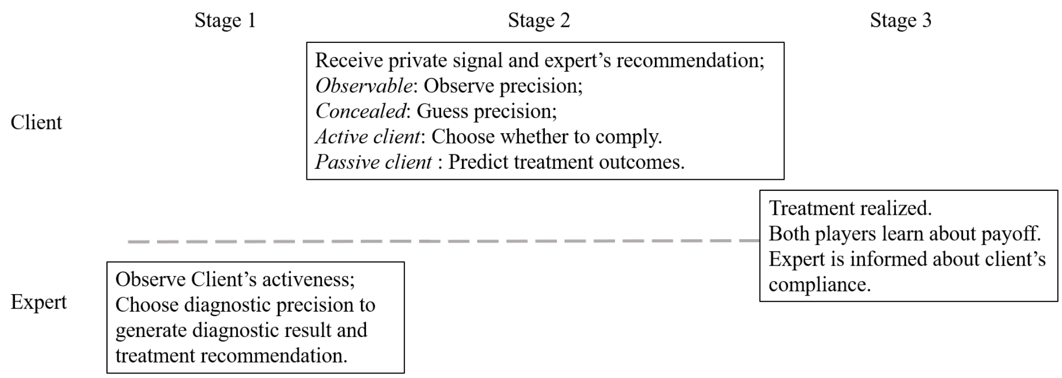


Figure 3: Observable Condition & Concealed Condition

probability of 30%, and this activeness status will be fixed for each client. I assigned more subjects to be active because the active client case is more important in this study.

To verify the propositions in the previous section, I vary the client's opportunity to observe diagnostic precision by imposing two treatments, the *Baseline (Observable condition)* and the *Concealed condition*, varying the opportunity for clients to observe experts' diagnostic precision levels.

Baseline (Observable Condition): The procedure in the baseline is shown in Figure 3. In Stage 1, the expert observes the client's activeness and chooses the (costly) precision level to generate a diagnostic result. The chosen precision level and the treatment recommendation from the diagnostic result both will be presented to the client in Stage 2, as well as the client's own private signal about the problem. If the client is active, then the client chooses whether or not to comply with the doctor. Therefore, the client can choose the treatment *opposite* to expert's recommended one if he/she likes. If the client is passive, the client does not do anything by the theory, but in the experiment, I asked these passive clients to guess whether or not the problem will be solved and a correct guess returns them 10 tokens. In Stage 3, the treatment is realized.

Concealed Condition: As shown in Figure 3, the game procedure of the Concealed condition differs from the Observable condition in that clients are not able to observe their matched experts' diagnostic precision levels. Instead, I elicit the client's belief about the precision level, which will help verify Proposition 3 that clients' compliance largely depends on their own beliefs about diagnostic precision. A correct guess of the precision rewards the client 10 tokens [7].

Furthermore, I impose two interventions, the *Communication condition* and the *Reputation condition*, which are both based on the setup of the Concealed condition.

[7]Clients do not know the correctness of their guess until the end of the whole experiment, to avoid the feedback of experts' precision affecting their subsequent beliefs and behaviors.

Communication: Similar to the Concealed condition, under the Communication treatment, clients cannot observe experts' diagnostic precision. However, as shown in Figure 4, subjects will enter a "negotiation" stage (highlighted in gray color) after experts choose diagnostic precision and before the diagnostic result is generated. In this stage, the matched client and expert will have an opportunity to chat, and one should expect clients to persuade experts to increase the diagnostic precision level or ask for information about the diagnostic precision.[8] After the negotiation stage, experts will have the opportunity to revise their diagnostic precision, then conduct the diagnosis and generate the treatment recommendation.

Reputation: Under the Reputation condition, clients still do not observe experts' diagnostic precision. There is a rating system for clients to rate their matched experts. As shown in Figure 5, at the beginning of each round, clients will read their matched experts' average ratings (on a scale of 1 to 5), and at the end of each round, clients will give ratings to their matched experts, which will be included into the calculation of the expert's average ratings for the subsequent rounds. Experts will be informed about the new rating given by the patient and the updated average rating. Hence, as the game continues, more and more ratings about each expert will be accumulated.[9]

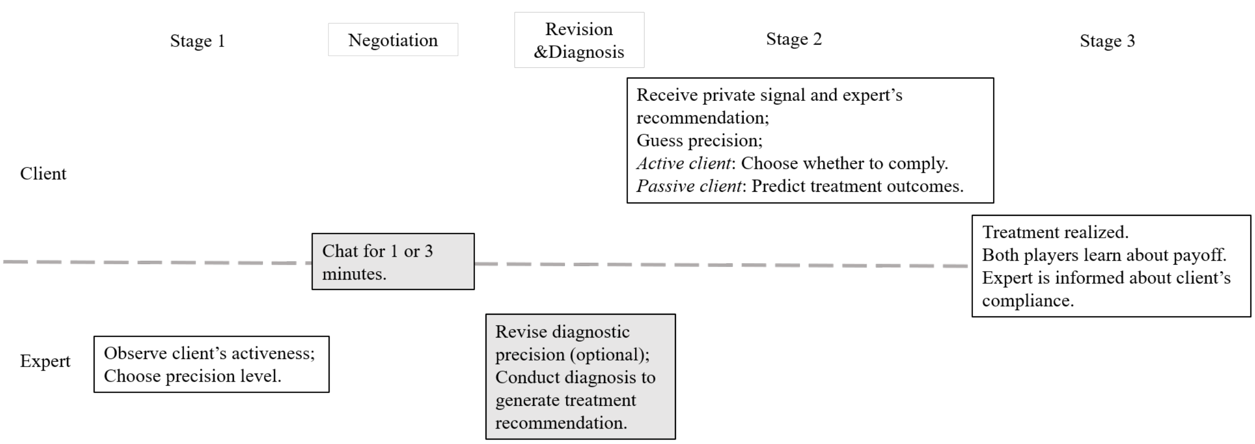


Figure 4: Communication Condition

[8]In the first round, they chat for 3 minutes. The time is reduced to 60 seconds for subsequent rounds, since they learn to communicate efficiently. Based on pilot sessions, 60 seconds is enough for their chatting. In the instruction, subjects are cleared informed that any information revealing their personal identity is be prohibited in the chat.

[9]In round 1, experts do not yet have any ratings. At the end of round 1, patients will give ratings to experts. Hence, starting from round 2, patients will observe the actual rating of their matched experts.

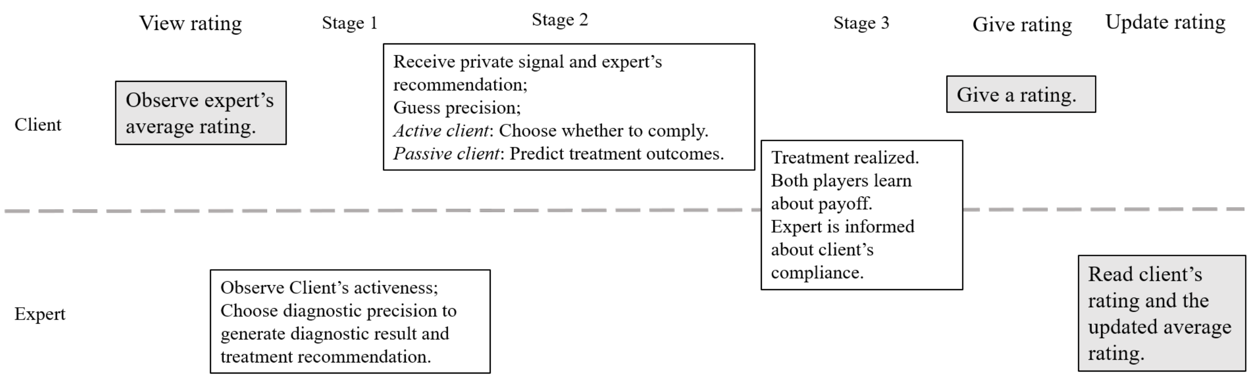


Figure 5: Reputation Condition

4.3 Experimental Procedure

Following previous practice (Fiedler et al., 2013), I separate preference measures from the main experiment to reduce spillover effects. Prior to participating in the lab experiment, all subjects were required to complete an online survey which consists two widely-used, incentivized tasks to measure risk aversion and altruism. The altruism measure consists of the dictator game suggested by Forsythe et al. (1994), implemented with a "role uncertainty" matching approach, under which all subjects make resource allocation decisions. They then are randomly matched into pairs with one member of the pair randomly chosen as the dictator whose decision will be implemented[10]. The risk preference is measured by the investment task suggested by Charness and Gneezy (2010). Subjects choose how much of a fixed endowment to invest in a risky asset, which gives a continuous measure of risk preferences. This measure is desirable because of its simplicity and ease of use with a variety of populations (Charness et al., 2013).

Each experimental session is assigned to one of the four treatment conditions discussed before. Once subjects arrive in the lab, they are randomly assigned to the role of either the expert or the client. Within each role, clients are randomly assigned to be "passive" or "active", and experts are randomly assigned with the alignment parameter γ. Their roles remain stable throughout the experiment. The matching of experts and clients is random for each round, avoiding the construction of reputation through repeated interaction. To further reduce any interdependency between rounds while retaining incentive compatibility, subjects are informed at the beginning of the experiment that one out of 20 rounds will be randomly selected for payment (see Charness et al. (2013) and Azrieli et al. (2018) for discussion on

[10] Iriberri and Rey-Biel (2011) pointed out that the use of role uncertainty may overestimate the prevalence of social-welfare-maximizing preferences within the sample. However, since I am interested in the correlation between altruism and doctors' choices in diagnostic precision rather than the overall level social preferences, this method provides credible measures in my context.

the incentive compatibility of the random payment scheme). At the end of the experiment, subjects fill out a questionnaire collecting their demographic information, their perceptions about the game, and their self-reported altruism and risk attitudes as a robustness check of the preference measure from the online surveys.

4.4 Predictions

In this subsection, I derive the predictions for the experiment, with a focus on the Observable and Concealed conditions.

4.4.1 Prediction of Diagnostic Precision with Passive Clients

Figure 6 characterizes the prediction of the expert's optimal choice of diagnostic precision when interacting with a passive client: $E^* = 0.5 + \frac{\gamma}{2}$. This means that the optimal diagnostic precision is increasing with the incentive alignment parameter γ. This prediction is a specific case of Proposition 1 derived from the numeric setup of the experiment. Notice that this prediction holds for both the Observable and Concealed conditions, because when the client is passive, whether or not the client observes the expert's diagnostic precision should not affect the expert's choice. Recall that experts are assigned one of the four values $\gamma \in \{0, 0.2, 0.6, 1\}$ which induce different levels of incentive alignments with clients. In Figure 6 the corresponding choice of diagnostic precision for each level of γ is highlighted by a circle, from $E^* = 0.5$ (the lowest precision, which is cost-free)to $E^* = 1$ (the highest precision that always generates correct diagnosis). In sum, prediction 1a below is the testable prediction for those round with passive clients:

Prediction 1a *When interacting with passive clients, experts' diagnostic precision is increasing with γ, from $E = 0.5$ to $E = 1$.*

4.4.2 Predictions of Diagnostic Precision with Active Clients

Consistent with the prediction from Lemma 1, under both Observable and Concealed conditions, I predict that in this experiment, clients will always follow their experts when their private signals are consistent with experts' diagnostic results. When an information conflict occurs, following Lemmas 2 and 3, an active client is predicted to go against the expert if she observes/believes that the diagnostic precision $E < 0.6$ under the Observable/Concealed condition. In the rest of this subsection, I will only discuss the case when there exists an information conflict between a client's private signal and an expert's diagnostic result.

To simplify the discussion, I will name experts with γ =0 or 0.2 as "low-alignment experts", and experts with γ =0.6 or 1 as "high-alignment experts", indicating different

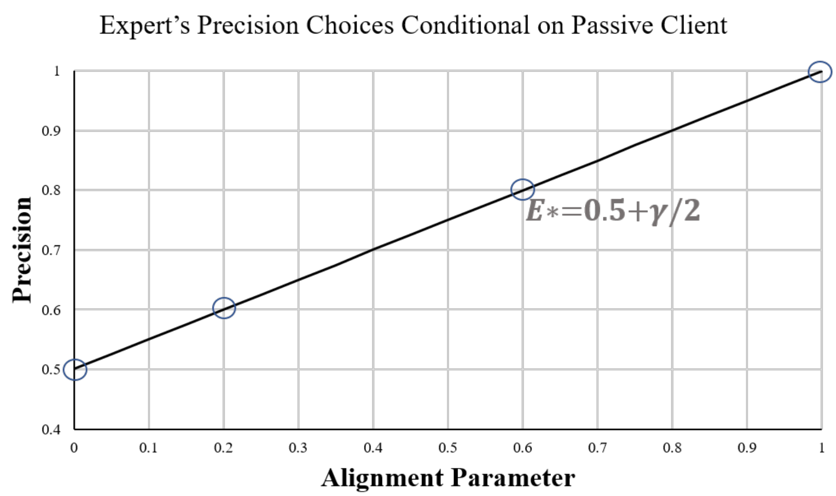


Figure 6: Expert's Diagnostic Precision with Passive Clients

Note. Circles in the figure are the predicted precision level choice for each assigned alignment parameter γ to experts.

levels of concerns that these experts have about clients' well-being. This is because experts with $\gamma < 0.4$ and experts with $\gamma \geq 0.4$ are behaving very differently, which will be explained below.[11]

Figure 7 depicts the equilibrium prediction of the expert's diagnostic precision and the client's compliance under the Observable condition. Low alignment experts are predicted to choose not to invest anything in improving diagnostic precision. On this equilibrium path, clients are predicted to never follow those experts. Notice that experts with $\gamma = 0.2$ are predicted to also choose the lowest precision of 0.5, which differs from the passive client case. Intuitively, this is because given that clients are active, experts are facing a tradeoff: they can either invest efforts to help clients improve their expected outcome, or choose not to exert any efforts and have clients rely on their private signals to solve their problems. For experts whose concern about the clients are not high enough ($\gamma < 4k(q-0.5)/(H-L)$), they will choose the latter strategy. In contrast, for high-alignment experts, they are predicted to invest to improve the diagnostic precision, following $E^* = 0.5 + \frac{\gamma}{2}$. Notice that these pairs of experts and clients return identical outcome as those in the passive client case. Therefore, in the experiment, one should predict identical diagnostic precision level between the active and passive clients under the Observable condition.

Prediction 2a *Conditional on Observable condition, for high-alignment experts (*$\gamma = 0.6$

[11] The threshold $\gamma = 4k(q-0.5)/(H-L)$ in Propositions 2 and 3 equals 0.4 based on the numerical setup of the experiment.

or 1*), their diagnostic precision for passive and active clients are the same.*

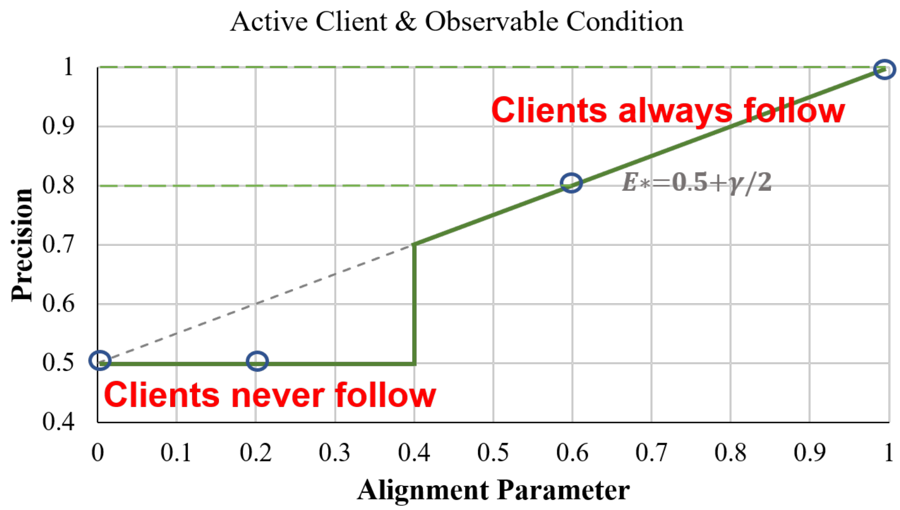


Figure 7: Equilibrium Prediction of Observable Condition

Note. Circles in the figure are the predicted diagnostic precision choices for each assigned alignment level γ to experts. The dashed line depicts the prediction of experts' diagnostic precision when clients are passive, as a benchmark for comparison.

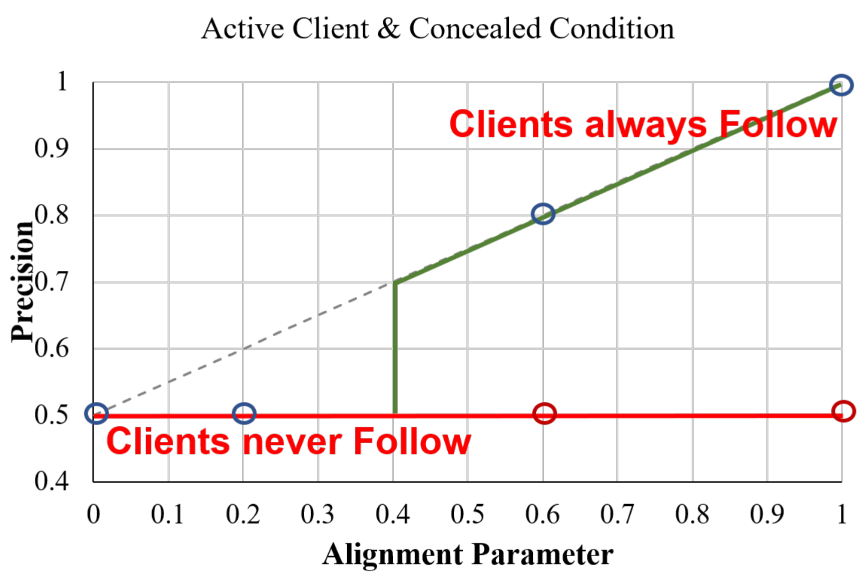


Figure 8: Equilibrium Prediction of Concealed Condition

Note. Circles in the figure are the predicted diagnostic precision choices for each assigned alignment level γ to experts. The dashed line depicts the prediction of experts' diagnostic precision when clients are passive, as a benchmark for comparison.

Figure 8 characterizes the equilibrium prediction under the Concealed condition. This equilibrium is derived from Proposition 3. Similar to the Observable condition, for low-alignment experts, under the Concealed condition, there exists only one equilibrium where clients believe that $\hat{E} < 0.6$. On this equilibrium path, experts choose the lowest diagnostic precision, and clients never follow their treatment recommendation. However, different from the Observable condition, for high-alignment experts, two equilibria coexist under the Concealed condition. In the first equilibrium, clients believe that $\hat{E} \geq 0.6$. On this equilibrium path, experts invest to achieve diagnostic precision of $E^* = 0.5 + \frac{\gamma}{2}$ and clients always follow experts. This equilibrium returns identical outcome as in the Observable condition and the passive client case. In another equilibrium, clients believe that $\hat{E} < 0.6$, where experts choose $E^* = 0.5$ (i.e. no effort and lowest precision) and clients never follow experts.

A lab experiment can examine the existence of the second equilibrium discussed above in two ways. Firstly, focusing on active clients, compared with the Observable condition, the existence of this additional equilibrium under the Concealed condition will reduce the average diagnostic precision among align-alignment experts (shown as Prediction 2b below). Secondly, conditional on the Concealed condition, compared with passive clients, this additional equilibrium will reduce the average diagnostic precision for active clients. In other words, passive clients will receive higher diagnostic precision than active clients from high-alignment experts, and this difference in diagnostic precision increases with experts' incentive alignment with clients (shown as Prediction 2c).

Prediction 2b *Conditional on active clients: for low-alignment experts ($\gamma = 0$ or 0.2), their diagnostic precision levels are the same across the Observable and Concealed conditions; for high-alignment experts ($\gamma = 0.6$ or 1), their diagnostic precision in the Concealed condition is lower than in the Observable condition.*

Prediction 2c *Conditional on the Concealed condition: for low-alignment experts ($\gamma = 0$ or 0.2), their diagnostic precision levels are the same across the Observable and Concealed conditions; for high-alignment experts ($\gamma = 0.6$ or 1), their diagnostic precision in the Concealed condition is lower than in the Observable condition.*

Finally, I also summarize the predictions for the behaviors of active clients below, which are secondary.

Prediction 3a *Clients follow their experts if their private signal is consistent with experts' diagnostic precision.*

Prediction 3b *If there exists a conflict between the client's private signal and the expert's diagnostic result, clients overrule experts if they observe (believe) that the experts' diagnostic precision is higher than 0.6 under the Observable (Concealed) condition.*

5 Results

I begin by providing the basic background information on the subjects in Table 2. A total of 436 undergraduate students from Texas A&M University participated in the experiment. They were randomly assigned to one of the four treatments discussed in the previous section. The average earning from lab sessions is \$20.03, including \$10 participation fee. In addition, subjects received \$1.90 for the online survey that they did before lab sessions.

Table 2: Summary Statistics of Subjects' Background Information

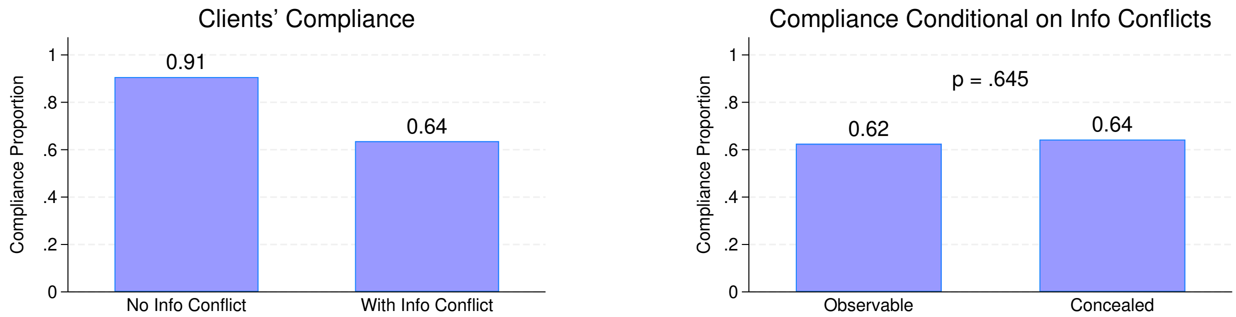
	Treatments				
	All	Observable	Concealed	Communication	Reputation
N. of Subjects	436	98	130	104	104
% Female	58.03	54.08	63.85	54.81	57.69
Age	20.11	19.71	19.91	20.41	20.42
	(1.85)	(1.48)	(1.47)	(1.83)	(2.44)
Risk Measure	49.80	53.95	52.13	46.04	46.74
	(26.70)	(27.11)	(28.47)	(25.06)	(25.10)
Altruism Measure	4.22	4.15	3.95	4.63	4.22
	(2.36)	(2.33)	(2.43)	(2.38)	(2.26)

Note: Standard deviations in parentheses. Risk measure is from the investment decision in Charness and Gneezy (2010), ranging from 0 to 100, with higher values indicating higher risk tolerance. Altruism measure is from the dictator game suggested by Forsythe et al. (1994), ranging from 1 to 11, with higher values indicating higher altruism level.

5.1 Active Clients' Participation in Treatment Selection

Before analyzing the impact of Active client's participation in treatment selection, it is important to first verify whether and how clients participate in selecting treatments.

Figure 9(a) depicts the proportion of compliance among Active clients by whether there exists information conflict between the client's private sinal and the expert's diagnostic result, pooling the Observable and Concealed conditions together. Consistent with Prediction 3a, when there is no information conflict, 91% of Active clients comply with their experts. However, when an information conflict occurs, this proportion reduces to 64%. Furthermore, Figure 9(b) restricts the data to those rounds with information conflicts and compares clients' compliance between the Concealed and the Observable conditions. This figure shows that



(a) Compliance by Info Conflicts

(b) Compliance with Info Conflicts

Figure 9: Active Clients' Compliance and Information Conflicts.

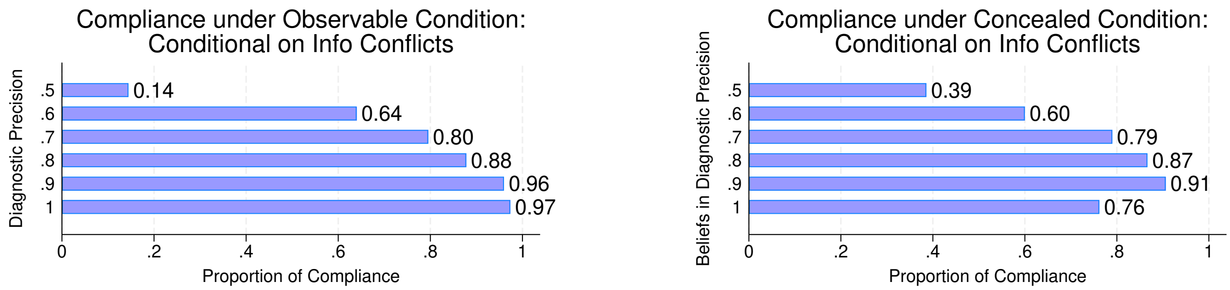
Note: "No Info Conflict' means the client's private signal is consistent with the expert's diagnostic result. "With Info Conflict" means the client's private signal differs from expert's the diagnosis.

compliance is not significantly different between these two treatment conditions.

Recall that Prediction 3b states that conditional on information conflicts, active clients' will overrule experts if the observed/believed diagnostic precision is lower than 0.6, which is the precision of their private signal. Figure 10(a) shows that under the Observable condition, when clients observe a diagnostic precision of 0.5, only 14% of clients comply with their matched experts, and this compliance proportion increases to 97% when the diagnostic precision rises to 1. Similarly, Figure 10(b) presents comparable information for the Concealed condition, where clients' compliance increases in accordance with their beliefs about the diagnostic precision.[12]

Overall, Active clients' choices in compliance are consistent with the theoretical prediction. More importantly, the findings above show that in the experiment, subjects assigned the role of Active clients are actively participating in decision-making – they go against their experts with valid reasons.

[12]One may notice that when the belief about precision $\hat{E} = 1$, only 76% of subjects comply with the expert, while they ought to fully follow the expert if they believe that the expert is 100% precise. I noticed from the data that there were several subjects who were not making consistent choices between belief elicitation and compliance: they always believed in the highest precision but chose to overrule the expert. For example, there is a subject with label 3009 who always chose $\hat{E} = 1$ but never followed. In Appendix Figure A1 I present again the figure by excluding this subject 3009. However, this does not mean that the result is not robust or a single subject can change the result, because subjects believe that $\hat{E} = 1$ for only 13.6% of time which makes it sensitive to only a single subject's choices.



(a) Observable Condition

(b) Concealed Condition

Figure 10: Proportion of Compliance Conditional on Information Conflicts

Note: Data is restricted to Active clients only. "Compliance" means that the active client chooses the treatment that the expert recommends. "Information Conflict" indicates those rounds where clients' private signal is not consistent with experts' diagnostic results.

5.2 Experts' Diagnostic Precision: When Does Clients' Activeness Matter?

This subsection analyzes the Observable and Concealed conditions to examine the predictions related to experts' behaviors, i.e., Prediction 1a, 2a, 2b, and 2c. The main variable of interest is *diagnostic precision* chosen by the expert, which also reflects the effort that the expert invests in achieving a specific diagnostic precision. In particular, we are interested in the following questions: How does the expert's level of concern for the client's well-being, as represented by the alignment parameter γ, influence their choices in diagnostic precision? Does the client's active role influence the expert's diagnostic precision? Furthermore, by comparing the Observable and Concealed conditions, I investigate the influence of concealing the diagnostic precision information on the expert's choices of diagnostic precision.

Table 3 presents the average diagnostic precision conditional on treatment conditions (Observable or Concealed) and experts' alignment parameter γ. Panel A is restricted to those rounds with passive clients, and Panel B focuses on those rounds with active clients. For both panels, to compare the diagnostic precision between the Observable and Concealed condition for each alignment parameter level, I run linear regressions on the diagnostic precision with standard errors clustered at the individual level as hypotheses testings, with the null hypothesis that there is no difference in the average diagnostic precision between the Observable and the Concealed conditions, controlling for repeated observations. These regressions employ a binary indicator that equals 1 under the Concealed condition and 0 under the Observable condition. The last rows of Panel A and Panel B of Table 3 report the p-value associated with the estimated coefficient of this binary indicator.

The key pattern from Panel A of Table 3 is that the average diagnostic precision is roughly increasing with experts' alignment parameters, for both Observable and Concealed conditions. This pattern supports Prediction 1a, indicating that when the expert has a higher concern about the client's well-being, then the expert will choose a higher diagnostic precision. Moreover, none of the p-values in Panel A achieves statistically significant, indicating the absence of difference in diagnostic precision between the Observable and Concealed conditions. This finding is in line with the theoretical prediction – intuitively, given that clients are fully compliant, whether clients observe the diagnostic precision does not change experts' choice of diagnostic precision.

However, one may notice that in Panel A of Table 3, there exists both over-investment among experts with the lowest alignment and under-investment among experts with the highest alignment, with more details as follows. When interacting with passive clients, the optimal choice of experts with the lowest alignment $\gamma = 0$ is $E^* = 0.5$, i.e. to not invest anything in the diagnosis. However, the observed average diagnostic precision conditional on $\gamma = 0$ is 0.58, which is significantly higher than the predicted level of 0.5 (p-value <0.001 for t-test of diagnostic precision = 0.5). One potential explanation is subjects' inherent altruism: there is a positive pairwise correlation of 0.1539 (p-value = 0.000) between experts' precision choices and their decisions in the dictator game, and this correlation increases to 0.5292 (p-value = 0.000) if restricting to experts with $\gamma = 0$. Therefore, subjects' higher altruism level prompts increased diagnostic precision even if they do not earn anything from solving the clients' problems. Furthermore, when $\gamma = 1$, the average diagnostic precision is 0.86 which is significantly lower than the optimal choice ($E^* = 1$). A possible explanation is subjects' loss aversion: there is a negative pairwise correlation of -0.1385 (p-value = 0.000) between experts' precision choices and their self-reported loss aversion, and this correlation remains -0.0556 (p-value = 0.0802) when focusing on experts with $\gamma = 1$.[13] Intuitively, if a subject is more loss-averse, they will avoid investing too heavily to prevent losses stemming from medical treatment failure.

Regarding active clients, Panel B of Table 3 reports the average diagnostic precision conditional on different incentive alignment levels. We observe that experts' diagnostic precision is increasing with the alignment parameter γ, i.e., experts with higher incentive alignment with clients are investing more in diagnostic precision. Moreover, none of the p-values in

[13]Experts' precision choice also exhibits correlation with risk tolerance captured by their decisions in the investment game conducted in the online survey. Their risk tolerance is negatively correlated with their precision choices. For instance, when restricting the sample to experts in the rounds of Observable and Concealed conditions interacting with passive clients, the correlation is -0.0514 (p-value = 0.0351). However, when further restricting the sample to experts with $\gamma = 1$, the correlation becomes 0.0085 and not statistically significant (p-value = 0.7896). Thus, risk tolerance does not appear to explain the under-investment of high-alignment experts in passive clients.

Table 3: Expert' Average Diagnostic Precision

Panel A: Diagnostic Precision for Passive Clients

	$\gamma = 0$	$\gamma = 0.2$	$\gamma = 0.6$	$\gamma = 1$	All
Both Conditions	.58 (.17)	.75 (.14)	.87 (.14)	.86 (.17)	.76 (.19)
Observable Condition	.57 (.14)	.76 (.13)	.85 (.13)	.83 (.17)	.74 (.18)
Concealed Condition	.61 (.20)	.73 (.15)	.87 (.14)	.89 (.15)	.79 (.19)
p-value from panel regression: Observable vs. Concealed	.29	.50	.86	.20	.23

Panel B: Diagnostic Precision for Active Clients

	$\gamma = 0$	$\gamma = 0.2$	$\gamma = 0.6$	$\gamma = 1$	All
Both Conditions	.57 (.14)	.67 (.15)	.77 (.15)	.79 (.18)	.70 (.18)
Observable	.53 (.08)	.72 (.14)	.80 (.14)	.83 (.18)	.71 (.18)
Concealed	.59 (.16)	.64 (.16)	.76 (.15)	.77 (.17)	.70 (.18)
p-value from panel regression: Observable vs. Concealed	.24	.13	.39	.65	.92

Note: Standard deviations in parentheses. p-values are from linear regressions using diagnostic precision as the outcome variable and the binary indicator that equals 1 for the Concealed condition and 0 for Observable condition as the independent variable, clustered at individual level. For each regression, I report the p-value associated with the estimated coefficient for the Concealed-condition indicator. These p-values indicate the statistical significance for the null hypothesis that the diagnostic precision is identical between the Observable and Concealed conditions, accounting for repeated observations from each subject.

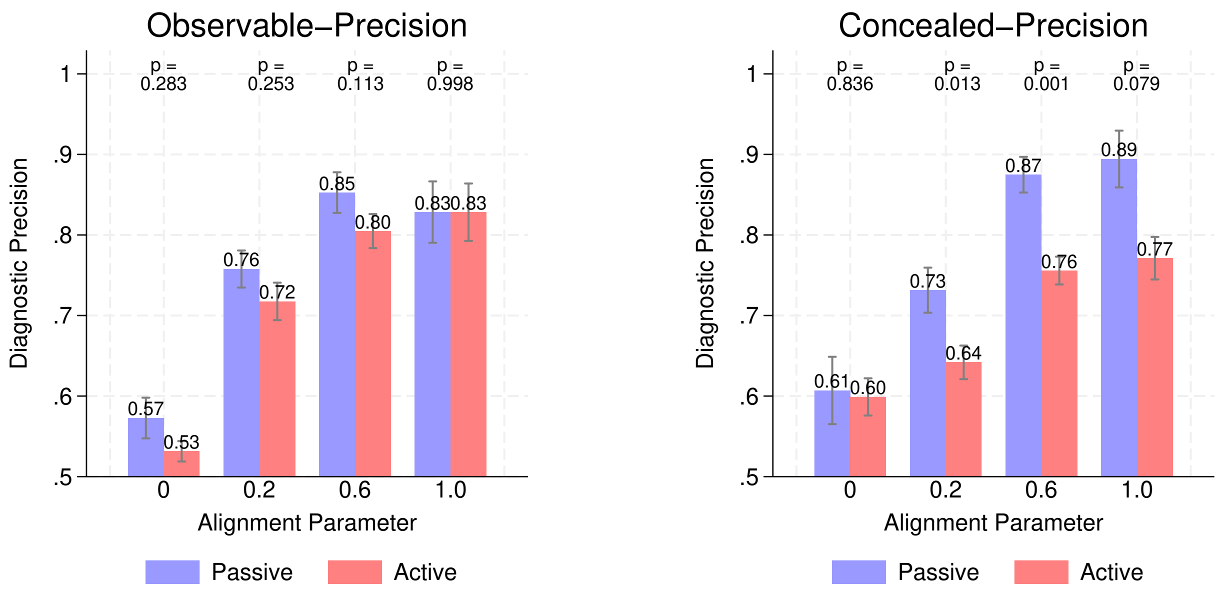
Panel B is statistically significant, suggesting that the average diagnostic precision between the Observable and Concealed conditions is not significantly different from each other, conditional on active clients. This finding appears not to align with Prediction 2b, which predicts that if the client is active, the average diagnostic precision under the Concealed condition should be lower than under the Observable condition. However, further exploration in the rest of this subsection reveals that the treatment effect of Concealed versus Observable conditions primarily influences experts' distinct approaches to active and passive clients, rather than exerting a direct impact on diagnostic precision.

To make direct comparisons of experts' attitudes toward passive vs. active clients, Figure 11 presents the average diagnostic precision by treatment conditions, clients' activeness, and experts' alignment parameter γ. An important difference between Figure 11(a) and 11(b) is the diagnostic precision gap between passive and active clients among high-alignment experts. In Figure 11(a), under the Observable condition, experts with $\gamma = 0.6$ or 1 do not choose different diagnostic precision between passive and active clients. This finding is in line with Prediction 2a that when the diagnostic precision is observable, high-alignment experts' diagnostic precision for active and passive clients are identical. In contrast, when the diagnostic precision is concealed from clients, as in Figure 11(b), high-alignment experts choose an average diagnostic precision level significantly higher for passive clients than for active clients. This finding supports Prediction 2c that when the diagnostic precision is not observable, high-alignment experts will treat passive and active clients differently – active clients are treated with lower diagnostic precision. This finding suggests that clients' activeness is discouraging experts from investing more efforts when the diagnostic precision is not observable.

To further verify findings from Figure 11 about the impact of clients' activeness on experts' diagnostic precision, Table 4 presents regressions with diagnostic precision as the dependent variable. The data is restricted to the Observable and Concealed conditions. Column (1) includes independent variables of the expert's alignment parameter, the matched client's activeness, the treatment conditions, and the interactions of these variables. Column (2) expands on this by including control variables such as round indicators, gender, ethnicity, altruism measure, risk attitude measure, whether the subject is of economics major, and whether the subject comes from Texas.

The regression analysis yields several findings. Firstly, the coefficient associated with the alignment parameter is statistically significant and positive, indicating that experts' diagnostic precision choice increases with their incentive alignment with clients. This increasing pattern is consistent with Predictions 1a. Secondly, the indicator of the client being active is associated with a significantly negative coefficient, suggesting that experts choose a lower diagnostic precision for active clients in comparison to passive clients. Notably, there is no evidence of a negative impact from the interaction of Concealed condition indicator and active-client indicator ($\beta = 0.005$ in Column (2) and not significant). Therefore, we do not find any direct effect of the Concealed condition in reducing diagnostic precision, which does not support Prediction 2b.

Due to the complexity of interpreting multiple interaction terms in Table 4, I derive the marginal effect of the Active indicator on diagnostic precision based on Column (2). Figure 12 visualizes the marginal effect of an active client across different treatment conditions



(a) Observable Condition

(b) Concealed Condition

Figure 11: Average Diagnostic Precision by Treatment Conditions, Alignment Parameter γ, and Clients' Activeness.

Note: Error bars are the 95% confidence intervals for averages. p-values are from linear regressions using diagnostic precision as the outcome variable and the binary indicator that equals 1 for Active clients and 0 for Passive clients as the independent variable, clustered at individual level. For each regression, I report the p-value associated with the estimated coefficient for the Active-client indicator. These p-values indicate the statistical significance for the null hypothesis that the diagnostic precision is identical between the Active and Passive clients, accounting for repeated observations for each subject.

and across different levels of alignment parameters, providing a clearer understanding of the results. This figure shows that for low-alignment experts ($\gamma = 0$ or 0.2), the negative impact of the client's activeness on the diagnostic precision is similar under the Concealed and Observable conditions. For high-alignment experts ($\gamma = 0.6$ or 1), the negative impact from the client's activeness is magnified under the Concealed condition, but is nearly eliminated to zero under the Observable condition. This finding is in line with the findings from Figure 11 that under the Concealed condition, the gap in investment in active and passive clients is enlarged when experts' incentive alignment with clients increases. This finding also provides evidence of the existence of the additional equilibrium that returns a "low-precision-low-compliance" outcome among high-alignment parameters discussed in Subsection 4.4.2.

Intuitively, this finding suggests that in an ideal scenario where a patient can observe how much effort a doctor exerts to achieve a certain diagnostic precision, it is straightforward for this patient to trust and follow the doctor. Therefore, an altruistic doctor will not worry

Table 4: Regression on Experts' Diagnostic Precision

	DV: Diagnostic Precision	
	(1)	(2)
Alignment Parameter	0.247***	0.206***
	(0.064)	(0.061)
Concealed	0.010	0.037
	(0.049)	(0.040)
Concealed × Alignment Parameter	0.054	0.054
	(0.095)	(0.087)
Active	-0.051*	-0.054**
	(0.026)	(0.021)
Active × Alignment Parameter	0.045	0.053
	(0.046)	(0.043)
Concealed × Active	0.009	0.005
	(0.040)	(0.033)
Concealed × Active × Alignment Parameter	-0.155*	-0.140*
	(0.084)	(0.074)
Constant	0.644***	0.653***
	(0.033)	(0.086)
Individual Controls	No	Yes
Round	No	Yes
Observations	2280	2280
Number of Individuals	114	114

Note: Individual controls include gender, ethnicity, whether the subject comes from Texas, subjects' altruism measured by choices in the dictator game, subjects' risk tolerance measured by choices in the investment game, and indicator of whether subjects major in economics or agricultural economics.
Standard error in parentheses, clustered at individual level. * $p < .1$; ** $p < .05$; *** $p < .01$.

about patients not listening to them. The doctor knows that the patient, after observing how much effort the doctor exerts to achieve a diagnosis that is more precise than the patient's self-diagnosis, will certainly be persuaded to follow the doctor's recommendation. On the contrary, in a more realistic setting where clients are unable to assess the diagnostic precision of a doctor, doctors with stronger concerns about patients' well-being may worry that patients will not treat their advice seriously. Patients' active participation in treatment selection creates a concern that they may overrule doctors' recommendation, making doctors' efforts in diagnosis a waste. Consequently, doctors with higher concerns for patients' health are more sensitive to whether a patient is fully compliant. Compared to patients who always fully follow doctors' advice, doctors give relatively less effort to patients who they think will not fully follow their advice.

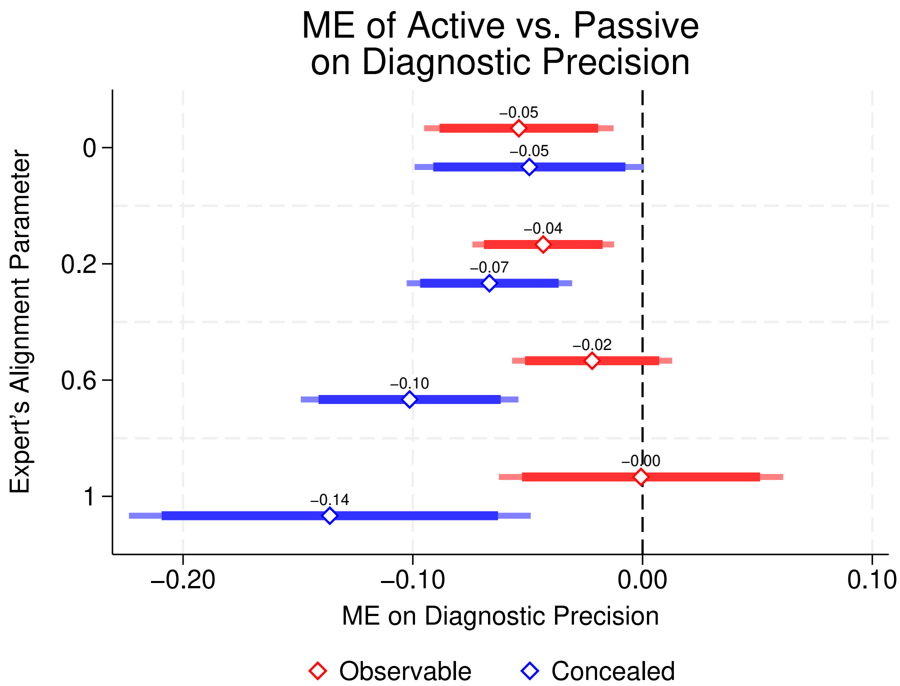


Figure 12: Marginal Effect of an Active Client by Treatment Conditions and Alignment Parameters γ

Note: This figure is derived from Model (2) in Table 4. The width of bars are with 95% and 90% confidence intervals.

5.3 Influence of Client Activeness on Client Welfare

Given the findings from the previous subsection that clients' activeness reduces experts' diagnostic precision under the Concealed condition among high-alignment experts, in this subsection, I will examine the influence of clients' activeness on their welfare, i.e. whether the problem is solved.

Table 5 summarizes the proportion of solved problems, divided by clients' activeness and treatment conditions. The proportion of solved problems is not statistically significantly different between Active and Passive clients under the Observable condition (p-value=0.233 from t-test). However, under the Concealed condition, Active clients solve significantly fewer problems than Passive clients (p-value < 0.001 from t-test). Moreover, regarding the comparison between Observable and Concealed conditions, Active clients solve fewer problems under the Concealed condition than under the Observable conditions, while Passive clients solve more problems under the Concealed condition. These comparisons indicate that clients' activeness plays an important role in predicting whether they can solve their problems.

Table 5: Proportion of Solved Problems

	Observable	Concealed	p-value
% Solved Problems (Active Clients)	69.11	61.48	0.003
% Solved Problems (Passive Clients)	72.62	82.14	0.001
p-value	0.233	<0.001	–

Note: p-values in the last row are from t-tests of the binary indicator of a solved problem, comparing Active vs. Passive clients; p-values in the last column are from t-tests comparing Observable and Concealed conditions.

Table 6: Marginal Effect of Clients' Activeness on Problem Solving

	DV: Indicator of Problem Solved	
	(1)	(2)
Active vs. Passive (Treatment = Observable)	-0.035	-0.056
	(0.034)	(0.034)
Active vs. Passive (Treatment = Concealed)	-0.207***	-0.207***
	(0.032)	(0.032)
Individual Controls	No	Yes
Round	No	Yes
Observations	2280	2280
Number of Individuals	114	114

Note: This table shows the marginal effect analysis of the effect of clients' activeness. The marginal effects are derived from logit regressions using the binary indicator of the problem solved (vs. not solved) as the dependent variable, with standard error clustered at individual level. Independent variables of the regression include the binary indicator of client's activeness (vs. passive), the treatment conditions (=1 if Concealed; =0 if Observable), and the interactions between these two variables. Individual controls include gender, ethnicity, whether the subject comes from Texas, subjects' altruism measured by choices in the dictator game, subjects' risk tolerance measured by choices in the investment game, indicator of whether subjects major in economics or agricultural economics.
Standard errors are in paratheses. * $p < .1$; ** $p < .05$; *** $p < .01$.

Table 7: Active Clients' Compliance and Welfare

	Observable	Concealed	p-value
% Mistaken Compliance	24.71	29.99	.0549
% Mistaken Nonompliance	51.97	71.04	.0006
% Solved Problems (Without Conflicts)	70.72	71.11	0.908
% Solved Problems (With Conflicts)	69.11	49.49	<.001

Note: p-values are from t-tests between the Observable and Concealed conditions; "% False Compliance" indicates the proportion of unsolved problems among those rounds where clients follow experts' advice; "% False Noncompliance" indicates the proportion of unsolved problems among those rounds where clients overrule experts' advice; "With Conflicts" indicates those rounds where clients' private signals are not consistent with experts' diagnostic results, while "Without Conflict" are those rounds without such information conflicts.

To further verify the findings above, I perform a logit regression using a binary indicator of a problem being solved as the dependent variable. Independent variables of the regression include the binary indicator of client's activeness (vs. passive), the treatment conditions (=1 if Concealed; =0 if Observable), and the interactions between these two variables. Then I perform marginal effect analysis of the effect of client's activeness, with results reported in Table 6.[14] Apparently, conditional on the Observable condition, being an Active client does not bring about any negative impacts on the probability of solving a problem. However, under the Concealed condition, the active role significantly reduces clients' probability of solving their problems.

There are two reasons why Active clients are less likely to solve problems, especially in the Concealed condition. The main reason is that high-alignment experts choose lower diagnostic precision for Active clients compared to Passive clients, which has been discussed in the previous subsection. Another reason is that Active clients under the Concealed condition are not able to correctly predict experts' diagnostic precision (see Figure A2 in the Appendix, which is a joint distribution of belief and actual diagnostic precision, reflecting that clients' beliefs are imprecise). The inability to predict experts' diagnostic precision increases the probability of making mistakes. Table 7 provides further evidence for the argument above. In this table, I define "Mistaken Compliance" as the case when a client follows an expert but fails to solve the problem, i.e., this client should have overruled the expert. Similarly, I define "Mistaken Noncompliance" as the case where a client overrules an expert where they should not have overruled. Under the Concealed condition, 29.99% of the rounds with compliance are Mistaken Compliance, and 71.04% of the rounds with noncompliance are Mistaken Noncompliance, which are both significantly higher than under the Observable condition (24.71% and 51.97% respectively). These differences indicate that there are more decision failures among Active clients under the Concealed conditions. As a result, the proportion of solved problems is lower under the Concealed condition than the Observable condition, especially when there exists information conflict between their private signals and experts' diagnostic results.

Bringing in the context of the patient-doctor relationship, the findings in this subsection demonstrate that if patients are not able to assess the precision of a doctor's diagnosis, then the patient's active involvement in treatment selection unintentionally reduces their well-being in the two ways. First, doctors reduce their efforts, leading to a lower probability of a correct diagnosis. Second, patients are not able to tell how credible a doctor's recommendation is, therefore their active involvement actually increases the risk of treatment failure.

[14] See Table A1 in the appendix for the raw logit regression result.

5.4 Effectiveness of Communication and Reputation Conditions

This subsection analyzes the effectiveness of the Communication and Reputation conditions in improving clients' welfare. The analysis specifically focuses on rounds involving Active clients.

5.4.1 Overview of the Impact of Communication and Reputation Conditions

Figure 13 compared the proportion of solved problems among Concealed, Communication, and Reputation conditions. This figure demonstrates that both the Communication and Reputation conditions lead to a significant increase in Active clients' proportion of solved problems compared to the Concealed condition (72% under Communication, 71% under Reputation, in comparison to 61% under Concealed). Furthermore, there is no significant difference in the proportion of solved problems between these two conditions. The positive impact of these two conditions on problem solving is further supported by a logit regression using the indicator of a problem solved as the dependent variable. Table 8 reports the marginal effects of the Communication and Reputation condition, using the Concealed condition as the baseline. Both the Communication and Reputation indicators are associated with positive and statistically significant marginal effects, suggesting that the Communication condition increases solved problems by about 10%, and the Reputation condition increases solved problems by around 9%. A coefficient test between these two treatment-condition indicators returns a p-value of 0.47 which suggests equal effectiveness of these two conditions in increasing the probability of solving a problem.

Why do these two conditions improve clients' well-being? In Figure 14, I overview both experts' and clients' behaviors. Figure 14(a) illustrates the average diagnostic precision for active clients, conditional on Concealed, Communication, and Reputation conditions. In both the Communication and Reputation conditions, the diagnostic precision is not significantly higher compared to the Concealed condition (0.727 for Communication, 0.715 for Reputation, versus 0.696 for Concealed) under clustered regression tests. Therefore, there is a lack of evidence that these two conditions improve clients' well-being by encouraging experts to invest more in the diagnosis.

Turning to the client's side, Figure 14(b) portrays the influence of these two institutional changes on clients' compliance, conditional on information conflicts between clients' private signals and experts' diagnoses. This figure indicates that the proportion of compliance increases from 64% under the Concealed condition to 72% under the Communication condition. Therefore, enabling clients and experts to engage in conversation significantly enhances clients' compliance with experts' recommendations. However, compliance does not

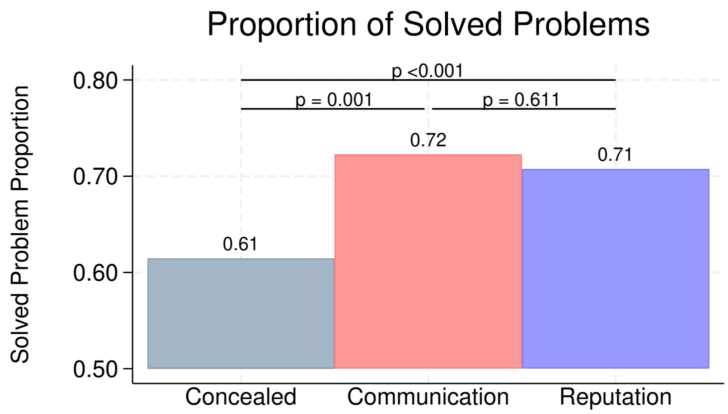


Figure 13: Proportion of problems solved among active clients, by treatment conditions

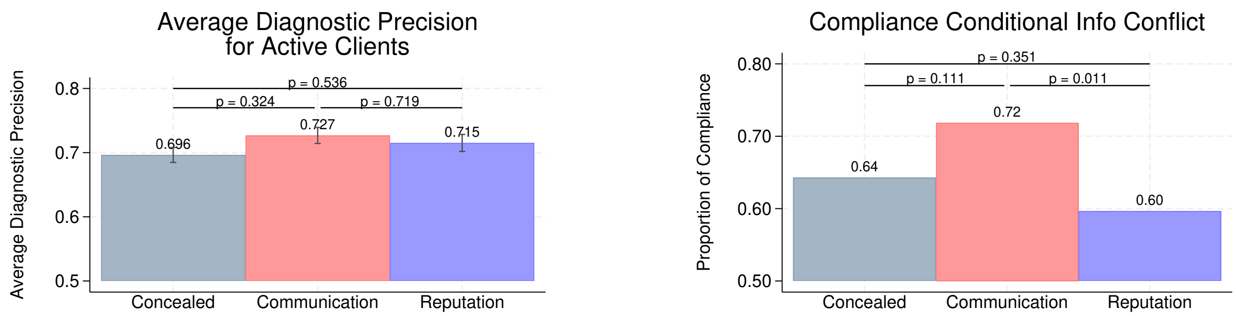
Note: p-values are from linear regressions using the indicator of solved problem as the dependent variable, clustered at individual level. Those p-values associated with the treatment condition indicators test the null hypothesis that there is no significant difference in solved problems between the two treatment conditions of interest, accounting for repeated observations for each subject.

Table 8: Marginal Effect of Communication and Reputation Conditions on Solved Problems among Active Clients

	DV: Indicator of Disease Solved	
	(1)	(2)
Communication	0.108***	0.108***
	(0.030)	(0.031)
Reputation	0.093***	0.086***
	(0.027)	(0.030)
Individual Controls	No	Yes
Round	No	Yes
Observations	2360	2360
Number of Individuals	114	114

Note: Regressions in this table are restricted to data of Active clients from the Concealed, Communication, and Reputation conditions, with the Concealed condition as the baseline. Both "Communication" and "Reputation" in the regression are binary indicators that equal 1 if the individual of a certain round is under this corresponding treatment condition. Individual controls include gender, ethnicity, whether the subject comes from Texas, subjects' altruism measured by choices in the dictator game, subjects' risk tolerance measured by choices in the investment game, indicator of whether subjects major in economics or agricultural economics.
Standard error in parentheses, clustered at individual level. * $p < .1$; ** $p < .05$; *** $p < .01$.

exhibit a significant increase under the Reputation condition (60%, in Figure 14(b))[15]. Hence, only the Communication condition effectively increases clients' compliance. The disparity in compliance between the Communication and Reputation conditions suggests that these institutional changes are likely to improve clients' well-being through distinct mechanisms, which will be discussed in the next subsections.



(a) Average diagnostic precision for active clients

(b) Proportion of compliance conditional on information conflict

Figure 14: Impact of Communication and Reputation Conditions on Experts' Precision Choices and Clients' Compliance

Note: Both panels restrict data to Active clients. "Info conflict" indicates those rounds when clients' private signal differs from experts' diagnosis. p-values in Panel (a) are from linear regressions using diagnostic precision as the dependent variable, clustered at individual level. Error bars in Panel (a) are the 95% confidence intervals for averages. p-values in Panel (b) are from linear regressions using the indicator of client following expert as the dependent variable, clustered at individual level. Those p-values associated with the treatment condition indicators test the null hypothesis that there is no significant difference in average diagnostic precision/client compliance between the two treatment conditions of interest, accounting for repeated observations for each subject.

5.4.2 Mechanism Investigation of Communication Condition

Under the Communication condition, experts have an opportunity to revise their diagnostic precision after chatting with Active clients. Experimental results show that experts revise the diagnostic precision only 14% of the time, leading to an average increase of 14.4 percentage points in diagnostic precision (from .680 to .824). However, as already mentioned in Subsection 5.4.1, this does not result in a significant increase in the average diagnostic precision. This is not a surprising finding, because the Communication condition does not change the incentive structure for experts.

[15]Although clients under the Reputation condition are less compliant, their decisions of whether to comply with experts are still consistent with their beliefs, i.e., they follow those experts if they believe that the diagnostic precision is higher than 0.6. See Figure A3 in the Appendix for clients' compliance by their beliefs under both Communication and Reputation conditions.

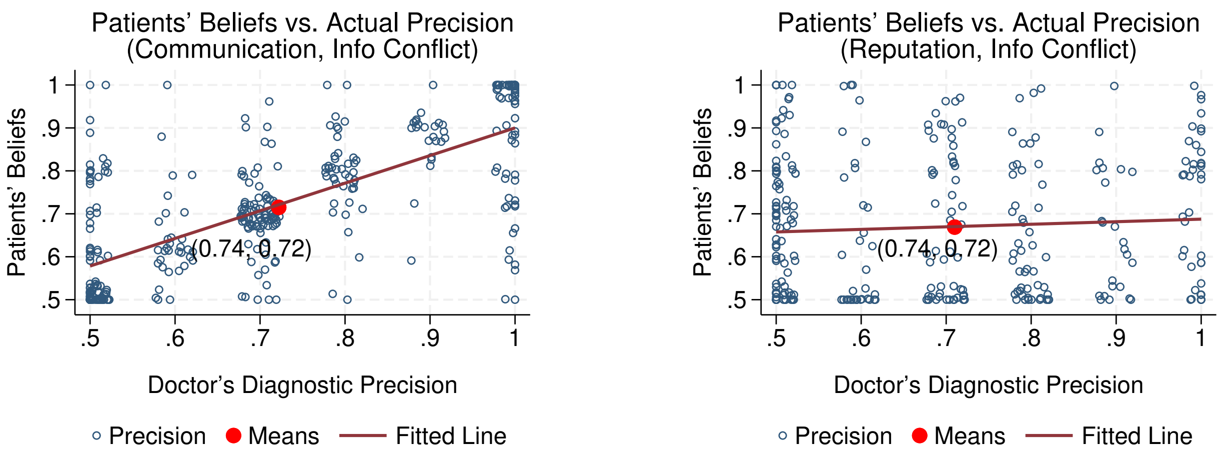
A more convincing reason behind the positive impact of Communication condition on clients' problem-solving is that communication enables clients to update their beliefs regarding experts' diagnostic precision and incentive alignment (see Figure A4 for the word cloud of chat messages, where "incentive alignment" and "diagnostic precision" are frequently mentioned in subjects' chats). Figure 15(a) depicts the distributions of clients' beliefs about diagnostic precision alongside the actual precision. In comparison to the Concealed condition (see Figure A2), under the Communication condition, clients' beliefs on the diagnostic precision become more accurate. This is evident from the upward-sloping fitted line with a statistically significant positive slope coefficient ($\beta = .643$, p-value $= .000$). Given that clients are making rational choices rewarding compliance based on their beliefs (as seen in Appendix Figure A3(a)), compared with the Concealed condition, clients under the Communication condition are better at following the highly precise medical treatment recommendations and overruling the imprecise ones, which drives the improvement of well-being as well.

5.4.3 Mechanism Investigation of Reputation Condition

While Figure 14(b) in Section 5.4.1 reveals that clients do not exhibit increased compliance under the Reputation condition compared to the Concealed condition, it is important to understand why clients still manage to improve their well-being. Figure 16 sheds light on the relationship between clients' compliance and experts' rating, and the relationship between experts' precision and the received rating.

In the left panel Figure 16(a), clients' observed expert average ratings are divided into 10 equal-width bins, and within each bin, the proportion of clients deciding to comply with their matched experts is plotted. This scatter plot demonstrates a positive correlation between the observed average rating and clients' compliance.This finding indicates that when facing an information conflict, clients are more inclined to follow experts who are of higher average ratings.

Moving to the right panel Figure 16(b), I further examine whether the average rating on experts conveys useful information about experts' actual diagnostic precision. Recall that after each interaction with experts, clients assign a rating to their experts, on a scale of 1 to 5 (5 the best). In Figure 16(b), for each rating level (1 through 5) that clients give, I calculate the corresponding average diagnostic precision. Notice that clients are unaware of their experts' diagnostic precision when assigning ratings. Surprisingly, the analysis reveals a positive association between ratings and diagnostic precision, indicating that clients assign higher ratings to experts with higher diagnostic precision. Therefore, clients' ratings effectively convey valuable information about experts' diagnostic precision, and as demonstrated in Figure 16(a), they adeptly utilize this information and follow high-rating experts.



(a) Communication Condition

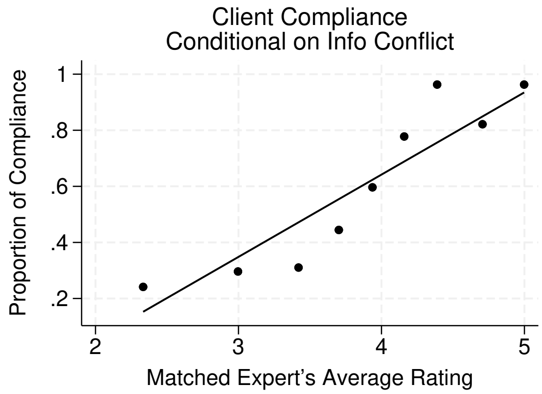
(b) Reputation Condition

Figure 15: Scatter Plots of Clients' Beliefs on Diagnostic Precision and Atual Diagnostic Precision

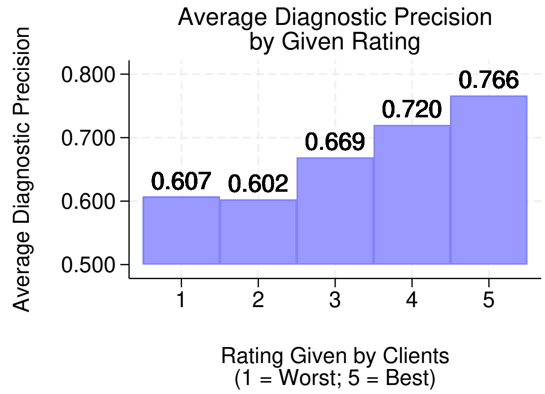
Note: Data for these two panels are restricted to rounds with Active clients where their private signals conflict with experts' diagnostic results. In both left and right panels, the red dot indicates the mean of the diagnostic precision and the mean of clients' believed precision; the fitted line is derived from the regression of the actual precision to the believed precision. On the left panel (Communication condition), the slope of the fitted line is $\beta = .643$ and p-value $= .000$. On the right panel (Reputation condition), the slope of the fitted line is $\beta = .060$ and p-value $= .285$.

In summary, compared with the Concealed condition, both the Reputation and Communication conditions increase clients' well-being by allowing them to solving more problems. Under the Communication condition, clients benefit from the ability to form more accurate beliefs about experts' diagnostic precision through conversations. This enables them to make informed decisions, following the recommendations of highly precise experts and disregarding those who are less precise. Moreover, the provision of precise information regarding experts' diagnostic precision greatly enhances clients' compliance. In contrast, the Reputation condition provides information about the experts in a less direct but still effective way. Rathering than directly learning about experts' diagnostic precision or incentive alignment, clients rely on a rating system to share and utilize valuable insights about their matched experts. Clients' ratings serve as an indirect indicator of diagnostic precision, allowing clients to follow recommendations from experts with higher ratings, thereby improving their well-being.

Taking medical services as the application, as discussed by Hanauer et al. (2014), the use of rating systems in gathering information about medical doctors is still not popular



(a) Active clients' compliance by experts' average rating conditional on info conflict. Experts' average ratings are divided into 10 bins. Conditional on each bin, I calculate the proportion of active clients who comply with those experts when there exists information conflicts.



(b) Average diagnostic precision by ratings given from clients after each round of interaction, conditional on information conflict.

Figure 16: Clients' Ratings, Experts' Diagnostic Precision, and Clients' Compliance

The left panel illustrates the relationship between experts' average rating and clients' compliance. The right panel illustrates the relationship between clients' ratings and experts' diagnostic precision.

compared with information gathering on restaurants, cars, movies, or books. Compared with training programs to improve the communication skills for patients and doctors, an online rating system is a less costly approach that still effectively increases patients' well-being. Therefore, I demonstrate the value of investing in a reliable rating system to provide effective guidance for patients when interacting with doctors.

6 Conclusion

This paper offers a comprehensive theoretical framework to discuss the influence of clients' participation in decision-making. I conduct a lab experiment to verify the predictions derived from the theoretical framework, which provides supporting evidence of the adverse effect of patients' active participation on doctors' efforts and patients' well-being.

The most important finding from this experiment is that when the client is not able to evaluate how much effort an expert exerts in improving the diagnostic precision, then the client's active participation in choosing treatments discourages experts from exerting more effort to achieve a more precise diagnosis. Intuitively, this is because compared with a paternalistic relationship, when clients have an option to go against experts, experts will consider their active participation as a possibility of being overruled. This discouraging effect is more pronounced among experts with stronger concerns about clients' problem solving. Furthermore, client's active participation in this scenario reduces their probability of solving their problems. One main reason is that experts already reduce their efforts which decreases the precision of diagnosis. Another reason is that given that clients are not able to assess different experts' diagnostic precision, their active participation is not beneficial, but rather even further reduces their probability of solving a problem.

To my best knowledge, those findings above have not been documented by empirical studies before, probably due to the challenges in observing experts' efforts and clients' participation in medical decision-making. By using a controlled experiment, I show a certain case where clients' active involvement in decision-making is making them worse off. However, one should notice that my finding does not suggest clients not participate in decision-making. Rather, I demonstrate the complex dynamics of expert-client interactions and point out a possible downside of clients' active involvement.

Furthermore, I investigate the effectiveness of two institutional changes, Communication and Reputation, in improving clients' problem solving. Both interventions demonstrated significant improvements in clients' outcomes, albeit through different mechanisms. Under the Communication condition, clients are able to engage in direct conversation with clients, leading to more accurate beliefs about experts' diagnostic precision. This enhanced information guided clients to follow highly precise treatment recommendations and disregard imprecise ones. On the other hand, the Reputation condition relies on a rating system to provide useful guidance for clients when interacting with experts. Patients' ratings on experts convey valuable information about experts' diagnostic precision, and they also effectively utilize experts' average ratings when making decisions regarding compliance with experts, ultimately leading to improved problem solving.

In the realm of healthcare services, the findings not only provide support for the importance of patient-doctor conversation, but also emphasize the value of investing in a rating system as a means of providing patients with valuable guidance when making compliance decisions. By utilizing such a system, patients can make more informed choices about their healthcare providers, leading to improved health outcomes. This highlights the importance of considering innovative approaches, such as rating systems, in the design of healthcare interventions. It would be valuable to collect field evidence for the impact of rating systems on healthcare providers' performance and patients' trust and adherence. I leave such questions for future research.

References

Alsan, M., O. Garrick, and G. Graziani (2019). Does diversity matter for health? experimental evidence from oakland. *American Economic Review 109*(12), 4071–4111.

An, P. G., L. B. Manwell, E. S. Williams, N. Laiteerapong, R. L. Brown, J. S. Rabatin, M. D. Schwartz, P. Lally, and M. Linzer (2013). Does a higher frequency of difficult patient encounters lead to lower quality care? *The Journal of family practice 62*(1), 24.

Azrieli, Y., C. P. Chambers, and P. J. Healy (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy 126*(4), 1472–1503.

Balafoutas, L., H. Fornwagner, R. Kerschbamer, M. Sutter, M. Tverdostup, et al. (2020). Diagnostic uncertainty and insurance in credence goods markets. Technical report.

Balafoutas, L. and R. Kerschbamer (2020). Credence goods in the literature: What the past fifteen years have taught us about fraud, incentives, and the role of institutions. *Journal of Behavioral and Experimental Finance 26*, 100285.

Banuri, S., D. De Walque, P. Keefer, and P. J. Robyn (2018). Encouraging service delivery to the poor: Does money talk when health workers are pro-poor? *World Bank Policy Research Working Paper* (8666).

Beresford, E. B. (1991). Uncertainty and the shaping of medical decisions. *Hastings Center Report 21*(4), 6–11.

Brekke, K. R., T. H. Holmås, K. Monstad, and O. R. Straume (2018). Socio-economic status and physicians' treatment decisions. *Health economics 27*(3), e77–e89.

Brosig-Koch, J., H. Hennig-Schmidt, N. Kairies-Schwarz, and D. Wiesen (2016). Using artefactual field and lab experiments to investigate how fee-for-service and capitation affect medical service provision. *Journal of Economic Behavior & Organization 131*, 17–23.

Brosig-Koch, J., H. Hennig-Schmidt, N. Kairies-Schwarz, and D. Wiesen (2017). The effects of introducing mixed payment systems for physicians: Experimental evidence. *Health Economics 26*(2), 243–262.

Charness, G. and U. Gneezy (2010). Portfolio choice and risk attitudes: An experiment. *Economic Inquiry 48*(1), 133–146.

Charness, G., U. Gneezy, and A. Imas (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization 87*, 43–51.

Darby, M. R. and E. Karni (1973). Free competition and the optimal amount of fraud. *The Journal of law and economics 16*(1), 67–88.

Dulleck, U. and R. Kerschbamer (2006). On doctors, mechanics, and computer specialists: The economics of credence goods. *Journal of Economic literature 44*(1), 5–42.

Dulleck, U. and R. Kerschbamer (2009). Experts vs. discounters: Consumer free-riding and experts withholding advice in markets for credence goods. *International Journal of Industrial Organization 27*(1), 15–23.

Emanuel, E. J. and L. L. Emanuel (1992). Four models of the physician-patient relationship. *Jama 267*(16), 2221–2226.

Fiedler, S., A. Glöckner, A. Nicklisch, and S. Dickert (2013). Social value orientation and information search in social dilemmas: An eye-tracking analysis. *Organizational behavior and human decision processes 120*(2), 272–284.

Fong, Y.-f., T. Liu, and D. J. Wright (2014). On the role of verifiability and commitment in credence goods markets. *International Journal of Industrial Organization 37*, 118–129.

Forsythe, R., J. L. Horowitz, N. E. Savin, and M. Sefton (1994). Fairness in simple bargaining experiments. *Games and Economic behavior 6*(3), 347–369.

Fox, S. and M. Duggan (2013). Health online 2013. *Health 2013*, 1–55.

Galizzi, M. M. and D. Wiesen (2018). Behavioral experiments in health economics.

Gottschalk, F., W. Mimra, and C. Waibel (2020). Health services as credence goods: A field experiment. *The Economic Journal 130*(629), 1346–1383.

Hanauer, D. A., K. Zheng, D. C. Singer, A. Gebremariam, and M. M. Davis (2014). Public awareness, perception, and use of online physician rating sites. *Jama 311*(7), 734–735.

Hennig-Schmidt, H., R. Selten, and D. Wiesen (2011). How payment systems affect physicians' provision behaviour—an experimental investigation. *Journal of health economics 30*(4), 637–646.

Hilger, N. G. (2016). Why don't people trust experts? *The Journal of Law and Economics 59*(2), 293–311.

Iriberri, N. and P. Rey-Biel (2011). The role of role uncertainty in modified dictator games. *Experimental Economics 14*(2), 160–180.

Kerschbamer, R., M. Sutter, and U. Dulleck (2017). How social preferences shape incentives in (experimental) markets for credence goods. *The Economic Journal 127*(600), 393–416.

Liu, F., A. Rasch, C. Waibel, and M. A. Schwarz (2020). The role of diagnostic ability in markets for expert services.

Lyu, H., T. Xu, D. Brotman, B. Mayer-Blackwell, M. Cooper, M. Daniel, E. C. Wick, V. Saini, S. Brownlee, and M. A. Makary (2017). Overtreatment in the united states. *PloS one 12*(9), e0181970.

Mimra, W., A. Rasch, and C. Waibel (2016). Second opinions in markets for expert services: Experimental evidence. *Journal of Economic Behavior & Organization 131*, 106–125.

Schulman, K. A., J. A. Berlin, W. Harless, J. F. Kerner, S. Sistrunk, B. J. Gersh, R. Dube, C. K. Taleghani, J. E. Burke, S. Williams, et al. (1999). The effect of race and sex on physicians' recommendations for cardiac catheterization. *New England Journal of Medicine 340*(8), 618–626.

Schulte, E. and M. Felgenhauer (2017). Preselection and expert advice. *International Journal of Game Theory 46*, 693–714.

Tan, S. S.-L. and N. Goonawardene (2017). Internet health information seeking and the patient-physician relationship: a systematic review. *Journal of medical Internet research 19*(1), e9.

A Additional Analysis

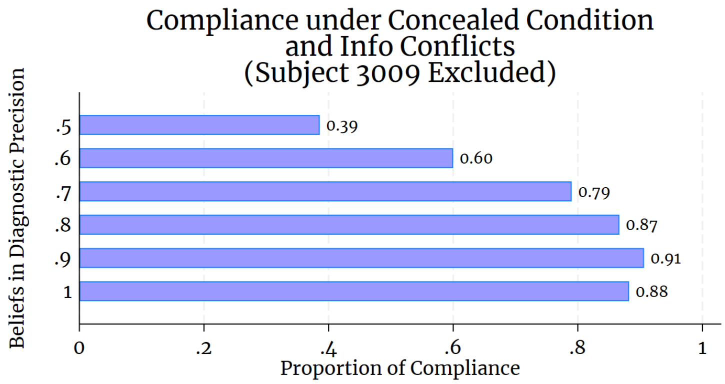


Figure A1: Patients' compliance under the Concealed condition when there exists information conflict, by patients' beliefs. In this figure, the Subject with label 3009 is excluded, as a comparison to Figure 9(b).

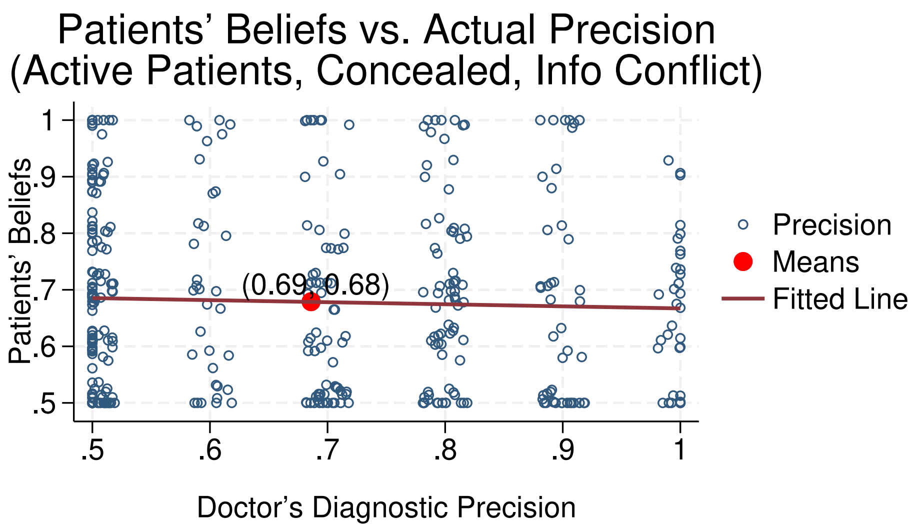
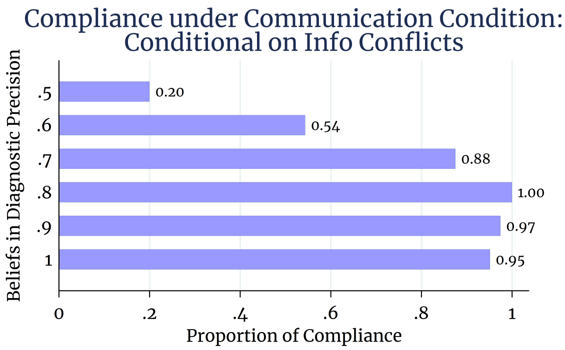


Figure A2: The jittered scatter plots of clients' beliefs on diagnostic precision and the actual diagnostic precision, conditional on active clients, Concealed condition, and information conflicts. The red dot indicates the mean of the diagnostic precision and the mean of client's believed precision. The fitted line is derived from the regression of the actual precision to the believed precision, with the slope $\beta = -.036$ and p-value $= .465$.

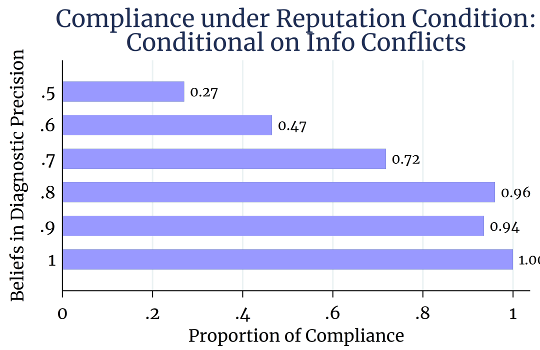
Table A1: Logit Regression: Influence of Client Activeness on Probability of Solving a Problem

	DV: Indicator of Prblem Solved	
	(1)	(2)
Concealed	0.551**	0.544***
	(0.218)	(0.210)
Active	-0.170	-0.270
	(0.167)	(0.167)
Concealed × Active	-0.888***	-0.807***
	(0.255)	(0.256)
Constant	0.975***	1.078***
	(0.131)	(0.281)
Individual Controls	No	Yes
Round	No	Yes
Observations	2280	2280
Number of Individuals	114	114

Note: Individual controls include gender, ethnicity, whether the subject comes from Texas, subjects' altruism measured by choices in the dictator game, subjects' risk tolerance measured by choices in the investment game, and indicator of whether subjects major in economics or agricultural economics.
Standard error in parentheses, clustered at individual level. * $p < .1$; ** $p < .05$; *** $p < .01$.



(a) Communication Condition



(b) Reputation Condition

Figure A3: Proportion of compliance conditional on information conflicts, by treatment conditions. Sample is restricted to active patients only. "Compliance" means that the active patient chooses the medical treatment consistent with the doctor's diagnostic result and treatment recommendation.

Figure A4: Word Cloud of Expert-Client Chat Messages in Communication Condition

Note: "Contract" is the framing of alignment parameter in the experiment; "accuracy" is the framing of diagnostic precision in the experiment.

B Lab Part Instructions

Introduction[16]

Thank you for participating in the experiment!

Please do not communicate with other participants during the experiment. If you have any questions, please raise your hand, and an experimenter will come to assist you. If you violate this rule, we may have to exclude you from the experiment and from all payments.

This experiment contains two parts. The first part is a decision-making game, which will repeat for 20 rounds, and at the end of the experiment, **one out of 20 rounds will be randomly selected and paid**. The second part is a questionnaire.

You will receive a $10 show up fee for your participation. This will be yours to keep. You also will have an opportunity to earn more, based on your decisions, the decisions of others, and luck. So, please read the following instructions carefully.

The experiment will use the currency named "tokens", with the following exchange rate:

1 token = 10 cents
(100 tokens = 10 dollars)

At the end of the experiment, your earnings in tokens will be converted into dollars. So, the more tokens you earn from the experiment, the more money you will be able to make. Your final payment will consist of three parts: (1) $10 show up fee; (2) your earnings from the online survey that you have already completed; (3) the earnings from today's experiment.

Instruction

In this experiment, there are two roles in a game: **client** and **expert**. You will be randomly assigned to one of the two roles, and will remain the same role for the whole experiment. There will be 20 rounds of games, and in each round, the matching of client and expert is random – in each round you should expect interacting with a different person.

Generally speaking, this game is about solving the client's problem. For each round, the client is facing a problem. The problem will always be either <u>Problem A</u> or <u>Problem B</u>, with equal probability (50%). This means that for different rounds, the client can have different problems. Each problem has its own correct solution: Problem A can be solved by <u>Solution A</u>*; Problem B can be solved by <u>Solution B*</u>.

The client does not know which problem he/she is facing, nor the correct solution to the problem. So, he/she has to consult an expert to identify the problem. The expert is randomly matched to him/her by the computer. The expert conducts a costly "<u>diagnosis</u>" to try to identify the problem and provide a recommendation of a solution to the client. If the client receives the correct solution to the problem, the problem will be successfully solved.

Next, you will read the detailed description of the client's and the expert's roles.

[16]This document is the instruction for the Observable condition. Instructions for other conditions are following the same format as this document.

Client's Role

Suppose you have been assigned as a client.

At the beginning of each round, you will be facing either Problem A or Problem B, and your goal is to solve your problem. If you receive the correct solution ("Solution A*" for having "Problem A"; "Solution B*" for having "Problem B"), you will get **120 tokens** for this round. If you receive the incorrect solution ("Solution B*" for having "Problem A"; "Solution A*" for having "Problem B"), you will get **20 tokens** for this round.

You do not know exactly which problem it is. Instead, you will receive a signal about the problem. The signal will be either "Problem A" or "Problem B", appearing on the computer screen. You should be aware that this signal may NOT be true: it only correctly identifies your problem **60%** of the time.

For example, if you see that the signal on the screen is "Problem A", you should be aware that there is a 60% chance that you actually have "Problem A", and 40% chance that you have "Problem B".

Another information source for you to guess your problem is the expert. You are randomly matched with an expert for each round. Your expert will conduct a "diagnosis" to identify your problem and recommend a solution to you. The way that the expert diagnoses your problem is by choosing the "diagnostic accuracy", i.e., how accurately he/she can identify your problem, ranging from 50% to 100%. 50% means that the expert's diagnosis will be correct only 50% of the time. And 100% means that the expert's diagnosis will be correct 100% of the time.

For example, suppose that the problem you are actually facing is "Problem A". If the expert chooses 50% as the diagnostic accuracy, then the expert's diagnostic result will be "Problem A" for only 50% of the time, and for the other 50% of time, the diagnostic result will be "Problem B". If the expert chooses 100%, then the expert's diagnosis will be "Problem A" for 100% of the time, which means that the expert will always be correct. (The expert's decision process will be explained later).

You will be able to observe the diagnostic accuracy level chosen by the expert.

After diagnosing your problem, the expert will recommend a solution to you with the following rule: if the expert's diagnosis is "Problem A", the expert always recommends "Solution A*"; if the expert's diagnosis is "Problem B", the expert always recommends "Solution B*".

As a client, you will be randomly allocated to be "Passive" or "Active". If you are Passive, you must follow the expert's recommended solution. If you are Active, you will freely choose the actual solution after seeing the expert's recommended solution. In other words, if you are "Passive", then your matched expert will choose a solution for you; if you are "Active", you will choose the solution yourself.

Additionally, if you are assigned to be Passive, we will ask you to predict whether your matched expert will solve your problem. Your correct prediction in a round will bring you extra 10 tokens of reward, if that certain round is randomly chosen for payment at the end of the experiment.

You will have a 30% chance to be assigned as a "Passive" client, and 70% chance to be assigned as an "Active" client. At the beginning of the experiment, you will be informed about the "Passive" and "Active" assignment, and this status will remain the same for the whole experiment.

For each round, your income is:

$$\textbf{Your income in this round} = \begin{cases} \textbf{120 tokens} & \textbf{if problem is solved} \\ \textbf{20 tokens} & \textbf{if problem is not solved} \end{cases}$$

Expert's Role

Now suppose that you are the expert.

You will be matched with a client for each round. Your matched client does not know which problem (Problem A / Problem B) he/she has, and you have the ability to diagnose the problem. Your diagnosis of the problem will also determine a recommendation of a solution for the client (Solution A* or Solution B*).

To diagnose the client's problem, you will first choose your diagnostic accuracy, which determines the probability that you correctly identify the client's problem. The accuracy level is associated with costs in tokens, according to the following cost table. You should notice that the higher the diagnostic accuracy you choose, the more you have to pay for this diagnosis.

Diagnostic accuracy	50%	60%	70%	80%	90%	100%
Cost of diagnosis in tokens	0	1	4	9	16	25

After choosing the accuracy level, the computer will generate a diagnosis of either "Problem A" or "Problem B" with your chosen accuracy. As already explained, the higher the accuracy level you choose, the more likely you will be able to correctly identify the client's problem. Your diagnosis will automatically generate a corresponding recommendation and be sent to your client (Solution A* for Problem A; Solution B* for Problem B).

However, some clients might not necessarily follow your recommendation. If the client is "Active", then he/she will have an opportunity to freely choose the solution after seeing your diagnosis and recommendation, which means that he/she may go against your recommendation. If the client is "Passive", he/she will always follow your recommendation. **For each round, we will inform you your matched client's type ("Active" / "Passive") before you choosing your diagnostic accuracy.**

For each round, you will receive **80** tokens as your base income.

Additionally, as an expert, you will be randomly assigned to one of the four contracts: 0-Contract, 20-Contract, 60-Contract, or 100-Contract. These contracts determine the extra income you will earn from successfully solving the client's problem. The following table summarizes the extra income from each type of contract:

Contract	Your Extra Income If Problem Solved
0-Contract	0 Tokens
20-Contract	20 Tokens
60-Contract	60 Tokens
100-Contract	100 Tokens

At the beginning of the experiment, you will be informed about the assigned contract, i.e., the extra tokens you can earn from solving the client's problem, and this will remain the same for the whole experiment. Your matched client will NOT be informed about your contract, i.e., the extra income that you are earning from solving his/her problem.

For each round, your income is:

Income for one round = 80 tokens + Extra Income from Problem Solving − Cost of diagnosis

Instructions, Continued

Now, the computer will randomly assign your roles. Notice that once you are chosen as a client or an expert, you will remain in the same role for the whole experiment.

After receiving your role assignment, you will go through a review about the procedure for your role, and then answer several questions to make sure that you understand the game rules.

Again, your decisions will affect how much you and your matched player can earn from the experiment, so, please pay attention to the instructions. If you have any questions, please raise your hand, and an experimenter will come to assist you.