Efficient Liability in Expert Markets

Chen, Yongmin and Li, Jianpei and Zhang, Jin

Department of Economics, University of Colorado Boulder, USA, Department of Economics, University of International Business and Economics, China, Department of Economics, University of International Business and Economics, China

3 June 2018

Efficient Liability in Expert Markets[*]

Yongmin Chen[†] Jianpei Li[‡] Jin Zhang[§¶]

March 02, 2022

Abstract. When providing professional services, an expert may misbehave by either prescribing "wrong" treatment for consumer's problem or failing to exert proper effort to diagnose it. We show that under a range of liabilities the expert will recommend the appropriate treatment based on his private information if markups for alternative treatments are close enough; however, a well-designed liability rule is essential for also motivating efficient diagnosis effort. We further demonstrate that unfettered price competition between experts may undermine the efficient role of liability, whereas either a minimum-price constraint or an obligation-to-serve requirement can restore it.

Keywords: Credence goods, experience goods, experts, liability, diagnosis effort, undertreatment, overtreatment.

JEL Codes: D82, I18, K13, L23

---

[*]Manuscript received October 2020; revised October 2021.

[†]Department of Economics, University of Colorado Boulder, USA; yongmin.chen@colorado.edu.

[‡]Department of Economics, University of International Business and Economics, China; lijianpei@uibe.edu.cn.

[§]Department of Economics, University of International Business and Economics, China; jzhang@uibe.edu.cn.

[¶]We are grateful to three referees and the editor, Masaki Aoyagi, for comments and suggestions that have substantially improved the paper. For helpful discussions and comments, we also thank Heski Bar-Issac, Andy Daughety, Yuk-fai Fong, Xinyu Hua, Ting Liu, Albert Ma, David Martimort, Yaofu Ouyang, Mike Riordan, Marius Schwartz, Kathy Spier, Curt Taylor, participants of the 2017 China Meeting of the Econometric Society, the 2017 HKUST Conference on Law and Economics, the 2018 Triangle Microeconomics Conference at UNC-Chapel Hill, the 2018 UIBE Workshop on IO and Competition Policy, Conference on Mechanism and Institution Design 2020, the Webinar Series on Credence Goods and Expert Markets 2021, and of seminars at Beihang University, Fudan University, Renmin University, Rutgers University, University of Leicester, and University of Toronto. Li gratefully acknowledges support by National Social Science Foundation of China (project number 19ZDA110). Zhang gratefully acknowledges support by National Natural Science Foundation of China (project number 71973024).

1. Introduction

When providing professional services, an expert often has superior information about the appropriate "treatment" for a consumer's problem. An extensive literature has studied how to prevent the expert from prescribing the "wrong" treatment for financial gains, with one major insight being that the expert's incentive to "cheat" can be removed if the markups for alternative treatments are equalized (Emons, 1997; Dulleck and Kerschbamer, 2006). In practice, the expert may also need to exert costly effort to diagnose the consumer's problem. The issue is then more complex and less well understood, especially because the "equal-markup" condition could eliminate the expert's incentive to exert diagnostic effort. In this paper, we investigate the role of liability in disciplining the expert's behavior in a model with both adverse selection and moral hazard.[1] We demonstrate that a well-designed liability can lead to efficiency in both the treatment recommendation and the diagnostic effort by the expert. We further show when the market may fail to be efficient even under the optimal liability, and what can be done to restore efficiency.

We consider a model in which a consumer needs a treatment for a problem (e.g., a medical condition) from an expert (e.g., a physician). The problem is either minor or major, and there are two alternative treatments that are competitively provided in the market. Upon seeing the consumer, an expert may immediately learn which treatment is appropriate from his expertise, or can exert (additional) private effort to obtain this information. The expert may then either provide a treatment or decline to serve the consumer without receiving any payment. He may prescribe the wrong treatment—a major treatment for a minor problem (*overtreatment*) or a minor treatment for a major problem (*undertreatment*)—if doing so increases his payoff. The type of treatment provided by the expert is observed publicly but the outcome of treatment is verifiable only with some probability.

Our setup departs from the existing literature on expert markets in two significant ways. First, we consider a general service product that differs from a pure credence or a pure experience good, with each as a limiting case, and we study broadly the optimal (i.e., welfare-maximizing)

[1] In our context, *adverse selection* refers to the deliberate inappropriate treatment by the expert given his private information, while *moral hazard* refers to a slack in the expert's diagnostic effort.

liability in expert markets. The literature has often considered goods/services in expert markets as credence goods (e.g., Darby and Karni, 1973; Taylor, 1995; Emons, 1997, 2001; Fong, 2005; Alger and Salanie, 2006; Liu, 2011). It thus makes the key assumption that consumers do not learn the treatment outcome afterwards—particularly in the case of overtreatment—and naturally precludes the use of liability to motivate experts. However, there is abundant evidence that service outcomes in expert markets, including healthcare, financial services, car repair services, auditing and taxi rides, can sometimes be verified, either through the use of modern technology (e.g., video recording the treatment process and comparative big data analysis) or with the help of third-party experts. Moreover, professional liabilities targeted at curbing negligent and fraudulent behavior in expert markets are an inherent part of tort law in many countries.

Second, unlike the focus in the literature on adverse selection, we also consider moral hazard in the model. The literature on credence goods often assumes that diagnosis is perfect and costless; thus, the issue of diagnostic effort does not arise.[2] We believe a model of both adverse selection and moral hazard captures more realistic features of many expert markets.[3] For example, a patient with a bad cough and a fever may need only a minor treatment (home rest, possibly with some medication) or a major treatment that requires hospitalization, and the physician may need the incentive to exert diagnostic effort to determine which treatment is appropriate, in addition to the incentive for truthful information reporting.[4]

We find that the presence of liability relaxes the incentive constraint for the expert to reveal his private information truthfully. Consequently, the equal-markup condition is no longer necessary to solve the adverse selection problem. For a wide range of liabilities, the expert will recommend the appropriate treatment based on his private information if markups for alternative treatments are

[2]See, e.g., Pitchik and Schotter (1987), Wolinsky (1993), Fong (2005), Hyndman and Ozerturk (2011), Fong, Liu, and Wright (2014), Hilger (2016), Fong and Liu (2018), Jost et al. (2021), Liu et al. (2020), Liu and Ma (2021). (We discuss several notable exceptions later.) By contrast, in our model the expert may need to be motivated to exert costly diagnostic efforts.

[3]Bardey et al. (2020) analyzes a market for experience goods that also combines adverse selection and moral hazard. They study optimal regulation and to what extent competition can substitute for regulation to curb the distortions from these two problems. Different from them, our paper studies optimal liability in markets that also share features of a credence good.

[4]The expert service could also be to repair a consumer's car, to fix a client's malfunctioning air-conditioning system, to provide advice on a client's legal problem, or to improve the security of a client's computer network. In all these situations, the expert may need to be provided with incentives both to exert costly diagnostic efforts and to report the consumer's problem honestly.

sufficiently close, and the familiar result under the equal-markup condition emerges in equilibrium as a special case of our model under zero liability.

While there are many liability rules under which the expert will recommend the appropriate treatment given his private information, they generally do not provide the efficient incentive for the expert's diagnostic effort. We derive the necessary and sufficient condition for a liability rule to result in both honest recommendation and efficient diagnosis, which we also term as the first best. An optimal liability is efficient if it achieves the first best. An efficient liability, when it exists, specifies damage payments for verified losses from wrong treatments that will induce equilibrium prices under which (i) the markup for each treatment is equal to its expected liability cost and (ii) the expected markup for the two treatments is equal to the efficient critical value of the expert's diagnosis costs. Then, the expert will conduct the additional diagnosis if and only if its cost does not exceed its expected social benefit, and he will also choose the efficient treatment—the treatment that maximizes the expected total surplus—based on his information.

We demonstrate that an efficient liability exists when, for instance, the expected loss to the consumer from inappropriate treatment is sufficiently high. However, it may fail to exist. Inefficiency can arise in our model for three possible reasons: the expert prescribes the wrong treatment given his information, he chooses diagnostic effort inefficiently, or he declines to serve the consumer after seeing her.[5] To reduce the expert's information rents when he learns the consumer's problem without additional diagnosis, the prices may become too low to incentivize the expert to exert the efficient diagnostic effort or to be willing to serve the consumer when the diagnosis cost is too high. Consequently, prices that maximize consumer surplus under unfettered price competition between experts can undermine the efficient role of liability, causing socially deficient diagnostic effort and treatment. This problem is exacerbated when the consumer is (partially) compensated through liability for her loss from a "wrong" treatment, because the social cost of such a loss is then not fully borne by the consumer. Our analysis also characterizes the second-best liability when the fully-efficient outcome is not attainable, under which the expert chooses the appropriate treatment given his information (or prior belief) but his diagnostic effort

[5]We assume that welfare is always higher for the consumer to receive some treatment than not to have any treatment.

is below the efficient level.

We further show that an efficient liability always exists if either (i) markups are constrained to be above certain minimum levels, or (ii) the expert is obligated to serve after seeing the consumer. Each of these two (regulatory) constraints, when feasible, ensures that under an optimal liability the expert will both exert efficient diagnostic effort and provide the proper treatment based on his information, thereby restoring full efficiency when it is undermined by prices that maximize consumer surplus. Intuitively, the (proper) minimum-price constraint directly provides the incentive for implementing efficient diagnostic effort. On the other hand, the obligation to serve removes the option for the expert not to treat the consumer after seeing her, so that the expert will efficiently choose between incurring an additional diagnosis cost or treating the consumer based only on his prior belief. In each of these two cases, the optimal liability satisfies the necessary and sufficient condition for efficiency and, in the case with obligation to serve, also ensures the expert's willingness to participate in the market. Moreover, if the treatment prices are set by a monopoly expert, an efficient liability also exists for all parameter values, though the equilibrium prices would then lead to the lowest consumer surplus.

The economic analysis of liability goes back to the seminal contributions by Brown (1973) and Shavell (1980). In markets where consumers can detect and verify a product's failure, the literature has studied how product liabilities affect a firm's incentives to improve product safety ex ante and to provide ex post remedy for an unsafe product (e.g., Daughety and Reinganum, 1995, 2008; Spier, 2011; Hua, 2011; Chen and Hua, 2012). Shavell (2007) presents a survey on the analysis of liabilities for accidents. In the existing literature on credence goods,[6] the role of liability in motivating the efforts and honesty of experts is rather limited. In most contributions, either consumers are assumed to be unable to tell ex post whether a treatment is appropriate (in which case liability cannot be effective as an incentive mechanism); or consumers can learn ex post only whether a problem is resolved and the institution of liability, defined as "the necessity for a seller to provide a good of sufficient quality to meet the consumer's needs" (Balafoutas and Kerschbamer, 2020), only prevents the experts from providing insufficient services. In contrast, we take the

[6]See Dulleck and Kerschbamer (2006) for a review of the earlier literature and Balafoutas and Kerschbamer (2020) for more recent contributions.

broader view of legal liability in the tradition of Brown (1973) and Shavell (1980), recognizing that products in expert markets often share properties of both credence and experience goods, and we analyze the judicious design of expert liability in a model of probabilistic verification for the treatment outcome (which can also be overtreatment), in accordance with the observed features in many expert markets.

Although most papers in the credence literature assume perfect diagnosis at zero cost and focus on the adverse selection problem, there are several notable exceptions. Dulleck and Kerschbamer (2009) investigates the incentives of experts to exert costly diagnostic effort and to report truthfully when the experts face competition from discounters who cannot perform diagnosis. Bonroy et al. (2013) extends this analysis by considering risk-averse consumers. Bester and Dahm (2018) analyzes optimal contract design when payment can be made contingent on the consumer's report of her subjective evaluation of the treatment outcome in a combined model of adverse selection and moral hazard. Balafoutas et al. (2020) analyzes a model in which the experts can make costly investment to reduce diagnostic uncertainty. On the other hand, Pesendorfer and Wolinsky (2003) abstracts away from the adverse selection problem and focuses on the interaction between the expert's choice of costly diagnostic efforts and the consumers' incentive to seek second opinions. We complement these studies by analyzing optimal liabilities in a combined model of adverse selection and moral hazard.

Our paper has relevant policy implications. While our model applies to markets with expert services in general, its most prominent application is probably the healthcare market where physicians' incentives are regulated by medical malpractice liabilities (e.g., Danzon, 1991).[7] Studies suggest that 4 to 18 percent of patients seeking care in hospitals in the U.S. are victims of medical malpractice, which could cost between $17-29 billion per year (Arlen, 2013). Liability for medical malpractice has emerged to discipline physicians and protect patients, but its performance has been controversial, and studies on its optimal design are scarce.[8] Our analysis sheds light on

[7]Many studies have found that physicians respond to financial incentives in treatment choices, including Gruber et al. (1999) on cesarean deliveries, Dickstein (2016) on the choice of drugs that treat depression, and Coey (2015) on treatment choices in heart attack management.

[8]As important exceptions, Simon (1982) compares negligence rule with strict liability in the healthcare market; Arlen and MacLeod (2005) analyzes optimal liability when the physician invests in expertise and there may be inadequate treatment. The key conflict in both papers is a moral hazard problem. Demougin and Fluet (2006, 2008) analyze the optimal assignment of liabilities under different rules of proof in lawsuits, which is applicable to

this issue. In particular, our results suggest that malpractice liability is essential for motivating physicians to exert proper diagnostic efforts. The efficient liability level depends not only on the magnitude of the loss, but also on whether there is overtreatment or undertreatment, because their probabilities of detection often differ. Furthermore, the welfare-maximizing liability is sometimes punitive, (much) exceeding the patient's loss from a malpractice incident.

The rest of the paper proceeds as follows. Section 2 presents the model, where given a liability rule, the consumer sets treatment prices followed by the expert's choices. Section 3 describes the efficient benchmark (i.e., the first best), and characterizes market equilibrium under a given liability rule. We show that the problem of finding a vector of the equilibrium prices or—equivalently—markups for the two treatments can be converted into one of finding a value of the equilibrium expected markup, which greatly simplifies our analysis. Section 4 analyzes the optimal liability. While the set of liabilities to consider is potentially very large, we show that the search for a welfare-maximizing liability can be confined to a much smaller class of liabilities. This enables us to derive a necessary and sufficient condition for the existence of an efficient liability that implements the first best, to identify the efficient liabilities under certain parameter restrictions, and to provide parameter values under which an efficient liability may fail to exist (in which case we also characterize the second-best liability). Section 5 establishes the results under a minimum-price constraint and under the obligation-to-serve requirement. Section 6 discusses alternative price regimes, where we argue that price-setting by the consumer is equivalent to price-setting by competitive experts in our model, and we further show that an efficient liability always exists in our model if prices are set by a monopoly expert. Section 7 concludes. The appendix contains all proofs that are not included in the main text.

2. The Model

A consumer needs a treatment (T) from an expert for a problem that can be either major or minor, $t \in \{M, m\}$, where $\Pr(t = m) = \theta = 1 - \Pr(t = M)$ and $\theta \in (0, 1)$. The expert can provide either a major treatment T_M or a minor treatment T_m, which is appropriate respectively

the healthcare markets, but their focus is very different from ours.

if t is M or m. The consumer's gross utility from the treatment is

$$(1) \qquad v(T,t) = \begin{cases} 0 & \text{if } T = T_t \text{ for } t \in \{M, m\} \\ -z_u & \text{if } T = T_m \text{ for } t = M \\ -z_o & \text{if } T = T_M \text{ for } t = m \end{cases}.$$

Thus, the consumer's gross utility is normalized to zero if she receives the appropriate treatment for her problem. If her type is M but the treatment is T_m, *undertreatment* occurs and the consumer suffers a loss $z_u > 0$. On the other hand, *overtreatment* occurs when problem type m is treated with T_M, in which case the harm to the consumer is $z_o \geq 0$.[9] If the problem is not treated by the expert, the consumer suffers an expected loss in the (absolute) amount of x. We further assume that the consumer is able to verify her loss z_u or z_o with probability $\alpha_u \in (0,1]$ or $\alpha_o \in (0,1]$, respectively when undertreatment or overtreatment has occurred. This formulation allows us to analyze a full spectrum of possibilities concerning the verifiability of the treatment outcome, encompassing pure credence goods and verifiable experience goods as the limit cases. In particular, the case of $\alpha_o \to 0$ and $\alpha_u \to 0$ corresponds to pure credence goods for which a consumer is unable to know the outcome after the treatment, the case of $\alpha_o \to 0$ and $\alpha_u = 1$ corresponds to full verifiability on undertreatment but no verifiability on overtreatment, and the case of $\alpha_o = \alpha_u = 1$ corresponds to an experience good for which the treatment outcome is learned by the consumer and is also verifiable. Most goods and services in expert markets probably fall between pure credence and verifiable experience goods with intermediate values of α_o and α_u.[10]

Note that the way we define the consumer's utility also differs from that in the credence goods literature, where the harm from overtreatment is usually normalized to zero, and undertreatment leads to the same utility as no treatment. (See, e.g. Emons, 1997; Dulleck and Kerschbamer, 2006). We depart from this modeling approach by assuming that overtreatment also leads to a

[9]Our analysis and results would be essentially the same if we interpret z_u and z_o as the expected losses associated with undertreatment and overtreatment.

[10]The loss might be verified not directly by the consumer, but by third party experts or the legal discovery process. In healthcare markets, overtreatment cases may center on the medical necessity of a procedure. For example, Dignity Health paid $37 million in a case for improper and medically unnecessary hospital admissions (Modern Healthcare, Oct. 30, 2014).

harm for the consumer (but allowing $z_o = 0$ as a special case) and undertreatment may lead to a loss different from no treatment (but with the two being equal as a special case). By adopting this more general—and possibly also more realistic—setup, we wish to explicitly account for the increasing concern over the harm from overtreatment in practice (e.g., Brownlee, 2008; Buck, 2013, 2015).[11]

Treatments T_M and T_m cost the expert C_M and C_m, respectively, with $0 \leq C_m < C_M$, and we assume

$$\text{(2)} \qquad \text{(i)} \quad \max\{C_M + \theta z_o, C_m + (1-\theta) z_u\} \leq x, \quad \text{and} \quad \text{(ii)} \quad \theta z_o < (1-\theta) z_u.$$

Thus, when not knowing whether $t = M$ or m: (i) applying a major or minor treatment is more efficient than leaving the problem untreated, and (ii) there exist parameter values under which T_M is either more or less efficient than T_m.[12] The type of treatment provided to the consumer—e.g., whether a certain procedure is carried out—is assumed to be publicly observed. Thus, if the expert recommends treatment T_t, cost C_t must be incurred to implement the treatment, for $t = M, m$.

The expert is better informed about the nature of the consumer's problem and, if necessary, can exert extra effort to diagnose the problem. Specifically, we assume that upon seeing the consumer, with probability $\beta \in [0, 1)$ the expert is informed about the realization of t (i.e., whether $t = M$ or m), while with probability $1 - \beta$ he is not informed of t but privately learns the realization of k, his private cost of diagnostic effort to learn the realization of t,[13] and k follows a continuous probability distribution $F(k)$ with density $f(k)$ on support $[0,\ \bar{k}]$. We denote the expert's decision on whether to incur k—if he does not observe the realization of t upon seeing the consumer—by $e \in \{E, N\}$. If he chooses E by incurring k, the expert learns the realization

[11] Buck (2015) reported that John Dempsey Hospital was discovered in 2011 to administer chest combination CT scans at nearly 10 times the national average while health experts noted that combination scans do not provide more valuable information in comparison to a single CT scan in most of those situations. Excess combination scans expose patients to large doses of radiation which increases the risk of developing cancer at later stage.

[12] When t is not known, T_M is more efficient than T_m if and only if $C_M - C_m < (1-\theta)z_u - \theta z_o$, and the inequality can either hold or reverse under different values of C_M and $C_m < C_M$ if $\theta z_o < (1-\theta) z_u$.

[13] This effort is beyond the observable normal effort associated with seeing the consumer. The extra cost k may include the additional time the expert spends with the consumer, the effort to gather additional information or the effort to learn new developments in treatment technology.

of t, while if $e = N$ (i.e., incurring no k) the expert maintains his prior belief about t.[14] Whether the expert incurs the diagnosis cost is his private information.[15] The expert, whose objective is to maximize expected profit, can refuse to treat the consumer after seeing her (which we denote as R). Under R, the payoffs to the expert and the consumer are respectively zero and $-x$.[16]

The expert may be liable for a bad outcome that is a result of maltreatment. The liability rule specifies damage payments $D \equiv (D_o, D_u)$, so that the expert is required to pay $D_u \geq 0$ if it is verified that the consumer has received undertreatment with loss z_u, and he is required to pay $D_o \geq 0$ if it is verified that the consumer has received overtreatment with loss z_o.

For a given liability rule D, our model is a sequential-move game between the two players, the consumer and the expert, with the following timing:

1. The consumer sets prices (P_M, P_m) for treatments T_M and T_m to maximize her expected surplus.[17] The consumer then visits the expert with her problem.

2. Upon seeing the consumer, the expert either learns the realization of t or, without learning t, the realization of his private cost of diagnostic effort k. In the latter case, he can privately choose $e \in \{E, N\}$. He then chooses $\Gamma_t \in \{T_M,\ T_m,\ R\}$ for $t \in \{M, m\}$ if he knows t (with or without k), or chooses $\Gamma_N \in \{T_M,\ T_m,\ R\}$ if $e = N$. The game ends if the expert chooses R, and it proceeds to the next stage otherwise.

3. The treatment recommended by the expert is implemented and payment (P_M or P_m) is made.

[14] In Pesendorfer and Wolinsky (2003), the expert similarly chooses diagnostic effort, where high effort has cost $c > 0$ and leads to perfect information, whereas low effort has zero cost and leads to no information. They assume that the expert provides honest recommendation when he has information but chooses a random recommendation when he has no information, and hence the adverse selection problem does not arise. In our model, prices and liabilities affect the expert's truthful reporting incentives.

[15] Notice that there are potentially four dimensions of the expert's private information: (i) whether he learns the realization of t upon seeing the consumer, (ii) the realization of k, (iii) whether he incurs the diagnosis cost, and (iv) whether $t = M$ or m (with or without incurring k).

[16] Notice that, as the expected value for the consumer's outside option, $-x$ could reflect the possibility that the consumer may visit other experts, for which there could be costs associated with delay or other frictions.

[17] This assumption is a shortcut to solve an isomorphic game with perfectly competitive experts in which equilibrium prices maximize consumer surplus. Relatedly, Arlen and MacLeod (2005) analyzes a setting in which the patients are price setters while the physician market is fully competitive. In Section 6, we will discuss alternative pricing assumptions and provide results when prices are restricted to be above some minimum levels or set by a monopoly expert.

4. If a loss from treatment is *verified*, the expert compensates the consumer according to the liability rule D.

In this consumer-expert game under a given D, a strategy of the consumer is a pair (P_M, P_m); while a strategy of the expert specifies, upon seeing the consumer and for each (P_M, P_m): (i) in the state where the expert immediately learns t, his choice $\Gamma_t \in \{T_M,\ T_m,\ R\}$ for $t \in \{M, m\}$, and (ii) in the state where k is needed to learn t, his private choice of $e \in \{E,\ N\}$ and his subsequent choice Γ_t if $e = E$ and Γ_N if $e = N$. Following each pair (P_M, P_m), there is a subgame in which the expert is the only player. A subgame perfect Nash equilibrium under a given D, which we shall also call a market equilibrium, is a pair of strategies by the consumer and the expert that induce an optimal choice by each player at her/his every decision point, given the strategy of the other player. As is usual for sequential-move games, we will solve the equilibrium of our game through backward induction. Our main interest is then to find a liability rule that would be chosen by a social planner to maximize welfare in equilibrium and to investigate when it achieves full efficiency, where we measure welfare by the expected total surplus of the consumer and the expert.[18]

3. Market Equilibrium

In this section, we analyze the market equilibrium for any given liability rule D. First, in Subsection 3.1, we describe the efficient benchmark. Next, in Subsection 3.2, we examine the expert's optimal choice of treatment strategy under given prices. We then derive in Subsection 3.3 the consumer's optimal prices—in anticipation of the expert's optimal choices—and they are thus also equilibrium prices for the consumer-expert game under any D. Finally, a numerical example is provided in Subsection 3.4 to illustrate the equilibrium.

3.1 Efficient Benchmark

Suppose all information is public and the expert can be required to act efficiently in all possible situations. If the expert learns t upon seeing the consumer, it is clearly efficient for him to choose

[18] In practice, the liability may be set by a government agency, the legislators, or the court.

T_M if $t = M$ and T_m if $t = m$. So we focus on the case where the expert needs to incur k in order to learn t. The expert can then choose (N, T_M): implementing T_M without incurring diagnosis cost k; or (N, T_m): implementing T_m without incurring cost k; or E_T: choosing E followed by selecting T_t for $t \in \{M, m\}$. The total surplus of the expert and the consumer for each of these choices is

$$(3) \qquad W(N, T_M) = -\theta z_o - C_M; \quad W(N, T_m) = -(1-\theta) z_u - C_m; \quad W(E_T) = -k - \overline{C},$$

where $\overline{C} \equiv (1-\theta) C_M + \theta C_m$. By part (i) of assumption (2),

$$W(N, T_M) = -\theta z_o - C_M \geq -x, \qquad W(N, T_m) = -(1-\theta) z_u - C_m \geq -x,$$

and thus if the expert has no additional information about t beyond his prior belief, a treatment has higher welfare than no treatment. Therefore it cannot be efficient for the expert to choose R. Define $\Delta C \equiv C_M - C_m$. Then $W(N, T_M) \geq W(N, T_m)$ if and only if

$$(4) \qquad \Delta C \leq (1-\theta) z_u - \theta z_o \equiv \Delta C^*.$$

That is, if the expert must choose the treatment based on his prior belief about t, it is efficient to choose T_M if the cost difference between the two treatments is sufficiently small ($\Delta C \leq \Delta C^*$), and to choose T_m otherwise. Notice that ΔC^*, which is positive by part (ii) of assumption (2), is increasing in z_u and decreasing in z_o.

Incurring the diagnosis cost is efficient when $W(E_T) \geq \max\{W(N, T_M),\ W(N, T_m)\}$, which holds if and only if

$$(5) \qquad k \leq \min\{\theta(\Delta C + z_o),\ (1-\theta)(z_u - \Delta C)\} \equiv k^*.$$

We assume $k^* \in (0, \bar{k})$ throughout the paper to focus on the more interesting case where it may be efficient—but not always efficient—to incur k.

Lemma 1 summarizes the efficient benchmark.

Lemma 1 *If the expert learns t upon seeing the consumer, it is efficient for him to choose T_t for $t \in \{M, m\}$. Otherwise, it is efficient to choose (i) (N, T_M) if $k > k^*$ and $\Delta C \leq \Delta C^*$; (ii) (N, T_m) if $k > k^*$ and $\Delta C > \Delta C^*$; (iii) E_T if $k \leq k^*$.*

Thus, when additional diagnostic effort is required to learn t, the efficient decision by the expert depends straightforwardly on the realized value of k and on the value of ΔC relative to ΔC^*: When the diagnosis cost is sufficiently high, it is efficient to have T_M without incurring k if the cost of major treatment is low, while it is efficient to have T_m without incurring k if the cost of major treatment is large; when the diagnosis cost is sufficiently low, it is efficient to incur k and then choose the appropriate treatment. For convenience, we shall say that the expert provides the right treatment if he chooses T_t for $t \in \{M, m\}$ when he knows t, whether he learns t initially or after incurring k; and when the expert needs to decide whether to incur k, it is understood that he has not learned t initially.

3.2 Expert's Choice of Treatment Strategy

Without loss of generality, denote any pair of prices by $P_M = C_M + \Phi_M$ and $P_m = C_m + \Phi_m$, where $\Phi_M \geq 0$ and $\Phi_m \geq 0$ are the markups over costs for the expert if he provides treatments T_M and T_m, respectively. Each pair of prices or—equivalently (Φ_M, Φ_m)—posted by the consumer is followed by a choice of the expert.

Since $\Phi_t \geq 0$ for $t \in \{M, m\}$, the expert never refuses to treat the consumer (i.e., choosing R) if he knows the realization of t. Furthermore, if the expert knows the realization of t, either upon seeing the consumer or after incurring k, it would be optimal for him to choose T_t for $t \in \{M, m\}$ *if and only if*

$$\Phi_M \geq \Phi_m - \alpha_u D_u, \qquad \Phi_m \geq \Phi_M - \alpha_o D_o. \tag{6}$$

Our analysis will proceed under the presumption that (6) holds—so that the expert will choose the right treatment if he knows t—and we later confirm that this is indeed the case in equilibrium, where it is optimal for the consumer to choose a pair of prices that satisfy (6) in anticipation of the expert's optimal strategy.[19]

Notice that for (6) to hold, $\Phi_M = \Phi_m$ if $D_u = D_o = 0$. When there are positive liabilities—as we allow in this paper—$\Phi_M = \Phi_m$ is sufficient but no longer necessary for (6): as long as the markups for the two treatments are not too different, the expert will have the right incentive to recommend the appropriate treatment if he knows t.[20] Thus, the presence of malpractice liability relaxes the constraint on markups that is needed to ensure the honesty of the expert.

Given (6), the expert will choose T_t for $t \in \{M, m\}$ if he learns the realization of t. Hence, we can focus our analysis on the expert's choice between R and the following three options if he does not initially learn t: (i) (N, T_M); (ii) (N, T_m); and (iii) E_T. The expert's profit from R is always zero. For a realized diagnosis cost k, the expert's profits from each of the other three choices are, respectively:

$$\pi(N, T_M) = \Phi_M - \theta\alpha_o D_o, \qquad \pi(N, T_m) = \Phi_m - (1-\theta)\alpha_u D_u, \tag{7}$$

$$\pi(E_T) = \theta\Phi_m + (1-\theta)\Phi_M - k, \tag{8}$$

where $\theta\alpha_o D_o$ is the expert's expected liability payment to the consumer under (N, T_M), since overtreatment occurs with probability θ; and similarly $(1-\theta)\alpha_u D_u$ is the expert's expected liability payment to the consumer under (N, T_m). The expert will make his choice to maximize his expected payoff; when he has the same expected payoff from any two options, we assume that he will choose the option that is favorable to the consumer.

Following a pair of prices with markups $\Phi \equiv (\Phi_M, \Phi_m)$, the expert's optimal choice when he does not initially learn t is E_T if and only if

$$\pi(E_T) \geq \max\{0, \pi(N, T_M), \pi(N, T_m)\}, \tag{9}$$

[19] To find an optimal liability, it is without loss of generality to devote our attention to situations where (6) is satisfied. If (6) is violated, the expert will have the perverse incentive to choose the "wrong" treatment even when he knows t, which cannot maximize welfare.

[20] This observation generalizes the implication of Dulleck and Kerschbamer (2006, Lemma 3) that the equilibrium markup can be smaller for major treatment than for minor treatment when the expert is liable for fixing the consumer's problem. It is also related to the idea in Bardey et al. (2020) that to motivate a seller to collect information and provide truthful advice on a consumer's choice between two goods, the profits from both goods must lie within an implementability cone. However, in our environment, markups close to each other do not guarantee the exertion of diagnosis efforts, and liability is crucial for such efforts.

or, equivalently, if $k \leq \hat{k}(D, \Phi)$, where

$$\text{(10)} \quad \hat{k}(D, \Phi) \equiv \min\{\theta\Phi_m + (1-\theta)\Phi_M, \;\; \theta\left(\Phi_m - \Phi_M + \alpha_o D_o\right), \;\; (1-\theta)\left(\Phi_M - \Phi_m + \alpha_u D_u\right)\}.$$

If $\Phi_M \geq \theta\alpha_o D_o$ or $\Phi_m \geq (1-\theta)\alpha_u D_u$, we have $\pi(N, T_M) \geq 0$ or $\pi(N, T_m) \geq 0$, and hence the expert will never choose R because, based on his prior belief about t, either T_m or T_M generates a (weakly) higher expected profit than not serving the consumer. In this case, the threshold $\hat{k}(D, \Phi)$ in (10) is simplified to

$$\text{(11)} \qquad \hat{k}(D, \Phi) = \min\{\theta\left(\Phi_m - \Phi_M + \alpha_o D_o\right), \;\; (1-\theta)\left(\Phi_M - \Phi_m + \alpha_u D_u\right)\}.$$

However, if $\Phi_M < \theta\alpha_o D_o$ and $\Phi_m < (1-\theta)\alpha_u D_u$, the expert earns negative expected profit when choosing T_M or T_m based on prior belief about t; therefore, he will choose R if the realized diagnosis cost is large. In this case, the threshold $\hat{k}(D, \Phi)$ in (10) is simplified to

$$\text{(12)} \qquad \hat{k}(D, \Phi) = \theta\Phi_m + (1-\theta)\Phi_M \equiv \overline{\Phi},$$

where $\overline{\Phi}$ denotes the (*ex ante*) expected markup of the two treatments. Notice that given a pair of markups satisfying (6), the expert will provide the right treatment if he knows t, with or without incurring k. We summarize the above discussions about the expert's optimal strategies following markups (Φ_M, Φ_m) in the result below.

Lemma 2 *When the expert knows t, he always chooses T_t for $t \in \{M, m\}$. When the expert does not initially learn t:*

(i) if $\Phi_M \geq \theta\alpha_o D_o$ or $\Phi_m \geq (1-\theta)\alpha_u D_u$, then the expert chooses E_T if $k \leq \hat{k}(D, \Phi)$ and chooses (N, T_M) or (N, T_m) otherwise.

(ii) If $\Phi_M < \theta\alpha_o D_o$ and $\Phi_m < (1-\theta)\alpha_u D_u$, then the expert chooses E_T if $k \leq \overline{\Phi}$ and chooses R otherwise.

3.3 Consumer's Optimal Prices

We now turn to the consumer's optimal choice of prices or, equivalently, markups. If the consumer does not receive any treatment, her surplus is $-x$. Her surpluses from the expert's choices (N, T_M), (N, T_m), and E_T are respectively

$$S(N, T_M) = -(C_M + \Phi_M) + \theta(-z_o + \alpha_o D_o), \tag{13}$$

$$S(N, T_m) = -(C_m + \Phi_m) + (1-\theta)(-z_u + \alpha_u D_u), \tag{14}$$

$$S(E_T) = -(1-\theta)(C_M + \Phi_M) - \theta(C_m + \Phi_m) = -\overline{\Phi} - \overline{C}, \tag{15}$$

where, for example, under (N, T_M) the consumer pays price $P_M = C_M + \Phi_M$, and she obtains expected utility $-z_o + \alpha_o D_o$ if t is m and 0 if t is M, which occur with probabilities θ and $1-\theta$, respectively. Notice that the consumer's expected surplus is always $S(E_T)$ when the expert knows t and chooses the right treatment, whether he learns t initially or after incurring k.

Because liability D can affect the expert's treatment choice and the consumer's price choice, it is useful for our analysis to impose some restriction on D to rule out liabilities that will clearly create undesirable incentives. At a minimum, D is such that the consumer should (weakly) prefer the expert to choose treatment T_t for $t \in \{M, m\}$ rather than to blindly choose T_M or T_m.[21] We shall thus focus on liability rules under which (i) $S(E_T) \geq S(N, T_M)$ when $\pi(N, T_M) \geq \pi(N, T_m)$, where the second inequality means that the expert may indeed choose (N, T_M); and (ii) $S(E_T) \geq S(N, T_m)$ when $\pi(N, T_m) \geq \pi(N, T_M)$.

For (i) above to hold, because

$$\pi(N, T_M) - \pi(N, T_m) \geq 0 \Longleftrightarrow \Phi_M - \Phi_m \geq \theta\alpha_o D_o - (1-\theta)\alpha_u D_u,$$

we have $S(E_T) - S(N, T_M) = \theta[\Delta C + \Phi_M - \Phi_m - (-z_o + \alpha_o D_o)] \geq 0$ if

$$\Delta C + z_o - (1-\theta)(\alpha_o D_o + \alpha_u D_u) \geq 0.$$

[21] In the latter case, a wrong treatment may occur and allow the consumer to collect a liability payment with probabilities θ or $1-\theta$.

Similarly, we can ensure (ii) above to hold if

$$z_u - \Delta C - \theta\left(\alpha_o D_o + \alpha_u D_u\right) \geq 0.$$

To guarantee both (i) and (ii), we thus focus on liabilities that satisfy

$$\alpha_u D_u + \alpha_o D_o \leq \min\left\{\frac{\Delta C + z_o}{1-\theta},\ \frac{z_u - \Delta C}{\theta}\right\}. \tag{16}$$

This condition prevents the consumer from having the perverse incentive to prefer the expert to choose a wrong treatment in order to collect the liability payment. In what follows, we proceed under the assumption that D satisfies condition (16), and we will later confirm that the condition indeed holds under any welfare-maximizing liability.

Under (16), the consumer would like the expert to provide T_t for $t \in \{M, m\}$, and under (6) the expert will indeed do so if he knows t. However, the expert will still need incentives, possibly with higher prices, to incur k or not to choose R. Recall from Lemma 2 that when $\Phi_M \geq \theta\alpha_o D_o$ or $\Phi_m \geq (1-\theta)\alpha_u D_u$, the expert will always choose to serve the consumer regardless of k; he will incur k to learn t if $k \leq \hat{k}(D, \Phi)$, and he will choose between (N, T_M) and (N, T_m) if $k > \hat{k}(D, \Phi)$.

Notice that it cannot be optimal for the consumer to choose $\Phi_M > \theta\alpha_o D_o$. To see this, suppose $\Phi_M > \theta\alpha_o D_o$. If $\Phi_m > 0$, then from (7) and (11), the consumer can increase her payoff by switching to $\Phi_M - \varepsilon$ and $\Phi_m - \varepsilon$ for sufficiently small $\varepsilon > 0$, which reduces prices but changes neither the expert's choice between (N, T_M) and (N, T_m) nor $\hat{k}(D, \Phi)$. If $\Phi_m = 0$, then $\hat{k}(D, \Phi) = \theta(-\Phi_M + \alpha_o D_o)$, and slightly reducing Φ_M would increase the consumer's expected payoff because: (a) it would increase $\hat{k}(D, \Phi)$, which is beneficial to the consumer since $S\left(E_T\right) \geq S(N, T_M)$ from (16), (b) it would simultaneously decrease the consumer's expected payment, and (c) it would not change the expert's choice between (N, T_M) and (N, T_m). Therefore, any markups with $\Phi_M > \theta\alpha_o D_o$ cannot be optimal for the consumer, and the same logic applies to markups with $\Phi_m > (1-\theta)\alpha_u D_u$. We therefore have:

Lemma 3 *For a given D, the consumer will optimally choose either (i) or (ii) below:*

(i) $\Phi_M = \theta\alpha_o D_o$ and $\Phi_m \leq (1-\theta)\alpha_u D_u$ or $\Phi_m = (1-\theta)\alpha_u D_u$ and $\Phi_M \leq \theta\alpha_o D_o$;

(ii) $\Phi_M < \theta\alpha_o D_o$ *and* $\Phi_m < (1-\theta)\alpha_u D_u$.

In case (i) of Lemma 3, the expert will always provide a treatment, and from Lemma 2 and (11) he incurs k if and only if k does not exceed

$$\begin{aligned}\hat{k}(D,\Phi) &= \min\{\theta\left(\Phi_m - \Phi_M + \alpha_o D_o\right), (1-\theta)\left(\Phi_M - \Phi_m + \alpha_u D_u\right)\} \\ &= \begin{cases} \theta\left(\Phi_m - \Phi_M + \alpha_o D_o\right) & if \quad \Phi_M = \theta\alpha_o D_o \text{ and } \Phi_m \leq (1-\theta)\alpha_u D_u \\ (1-\theta)\left(\Phi_M - \Phi_m + \alpha_u D_u\right) & if \quad \Phi_M \leq \theta\alpha_o D_o \text{ and } \Phi_m = (1-\theta)\alpha_u D_u \end{cases} \\ &= (1-\theta)\Phi_M + \theta\Phi_m = \overline{\Phi}.\end{aligned}$$

In case (ii), from Lemma 2 the expert will provide the right treatment if he knows t. Otherwise, he will choose E_T if $k \leq \overline{\Phi}$ and R if $k > \overline{\Phi}$.

It follows that when the consumer chooses prices optimally, the critical value for the expert to choose diagnostic effort is uniquely determined by the expected markup $\overline{\Phi}$:

$$\hat{k}(D,\Phi) = \overline{\Phi}. \tag{17}$$

Because the consumer's optimal markups satisfy $\Phi_M \leq \theta\alpha_o D_o$ and $\Phi_m \leq (1-\theta)\alpha_u D_u$, in equilibrium the expected markup $\overline{\Phi}$ under a given D is equal to

$$\overline{\Phi}_D \equiv \theta(1-\theta)(\alpha_o D_o + \alpha_u D_u) \tag{18}$$

if and only if $\Phi_M = \theta\alpha_o D_o$ and $\Phi_m = (1-\theta)\alpha_u D_u$, and $\overline{\Phi}_D$ is the highest possible $\overline{\Phi}$ in equilibrium under a given D.

Therefore, to find the consumer's optimal markups, it is without loss of generality to confine our attention to (Φ_M, Φ_m) with $\overline{\Phi} \in \left[0, \overline{\Phi}_D\right]$. Moreover, from (6) the expert will always choose T_t for $t \in \{M, m\}$ if he knows t; and from (17) his decision to incur k depends only on the value of $\overline{\Phi}$. Thus, the individual values of Φ_M and Φ_m will affect the consumer's surplus—beyond their impact through $\overline{\Phi}$—only when the expert does not initially learn t and also chooses not to incur k; and the consumer chooses Φ_M and Φ_m to maximize her surplus in this case—denoted as

S_{NT}—for a given $\overline{\Phi}$. This enables us to derive the consumer's optimal (Φ_M, Φ_m) for any given $\overline{\Phi} \in \left[0, \overline{\Phi}_D\right]$ and to express her expected surplus as a function of $\overline{\Phi}$, which converts the consumer's problem of finding the optimal (Φ_M, Φ_m) into a simpler problem of finding the optimal $\overline{\Phi}$. We formalize this idea in Lemma 4 below.

Lemma 4 *Given any liability D, the consumer's expected surplus is a function of $\overline{\Phi} \in [0, \overline{\Phi}_D]$:*

$$S\left(\overline{\Phi}\right) = \left[\beta + (1-\beta)F(\overline{\Phi})\right] S\left(E_T\right) + (1-\beta)\left[1 - F(\overline{\Phi})\right] S_{NT}\left(\overline{\Phi}\right), \tag{19}$$

in which $S(E_T) = -\overline{\Phi} - \overline{C}$ and

(20)

$$S_{NT}\left(\overline{\Phi}\right) = \begin{cases} -C_M - \theta z_o \equiv S_{NM} & if \quad \frac{\overline{\Phi}}{\theta(1-\theta)} \geq \alpha_o D_o \text{ and } \Delta C \leq \Delta C^*, \text{ or } \alpha_o D_o \leq \frac{\overline{\Phi}}{\theta(1-\theta)} < \alpha_u D_u \\ -C_m - (1-\theta) z_u \equiv S_{Nm} & if \quad \frac{\overline{\Phi}}{\theta(1-\theta)} \geq \alpha_u D_u \text{ and } \Delta C > \Delta C^*, \text{ or } \alpha_u D_u \leq \frac{\overline{\Phi}}{\theta(1-\theta)} < \alpha_o D_o \\ -x & if \quad \frac{\overline{\Phi}}{\theta(1-\theta)} < \min\{\alpha_o D_o, \alpha_u D_u\} \end{cases}.$$

The first term on the right-hand side of (19) corresponds to the consumer surplus when the expert knows t (either initially or after incurring k), while the second term corresponds to her surplus when the expert will choose among (N, T_M), (N, T_m) and R, where S_{NM} and S_{Nm} are respectively the consumer's maximal surplus when the expert chooses (N, T_M) and (N, T_m). By part (i) of Assumption (2), $S_{NM} > -x$ and $S_{Nm} > -x$. Notice that $S\left(\overline{\Phi}\right)$ can have at most two discontinuous points: (i) at $\overline{\Phi} = \theta(1-\theta)\min\{\alpha_o D_o,\ \alpha_u D_u\}$ where $S_{NT}\left(\overline{\Phi}\right)$ possibly jumps up from $-x$ to S_{NM} or S_{Nm}, and (ii) at $\overline{\Phi} = \theta(1-\theta)\alpha_o D_o$ if $\alpha_o D_o > \alpha_u D_u$ and $\Delta C < \Delta C^*$ or at $\overline{\Phi} = \theta(1-\theta)\alpha_u D_u$ if $\alpha_o D_o < \alpha_u D_u$ and $\Delta C > \Delta C^*$, where $S_{NT}\left(\overline{\Phi}\right)$ jumps up from S_{Nm} to S_{NM} or from S_{NM} to S_{Nm}. Therefore $S\left(\overline{\Phi}\right)$ is upper-semi continuous on $\left[0, \overline{\Phi}_D\right]$, and there exists

$$\overline{\Phi}^* \in \arg\max_{\overline{\Phi} \in \left[0, \overline{\Phi}_D\right]} S\left(\overline{\Phi}\right). \tag{21}$$

Combining Lemma 2, Lemma 4, equation (17), and equation (21), we have the following characterization of the market equilibrium.

Proposition 1 *In any equilibrium, the expert's threshold in his choice of k satisfies (17), and the consumer's choice of markups (Φ_M, Φ_m) satisfies (21). For any D, there exists an equilibrium, and the equilibrium is unique if $\overline{\Phi}^*$ is unique.*

For any D and for a given $\overline{\Phi}$, equation (33) in the proof of Lemma 4 in the appendix characterizes the optimal markups (Φ_M, Φ_m) that maximize S_{NT}, and the equilibrium markups (Φ_M^*, Φ_m^*) are then determined by setting $\overline{\Phi} = \overline{\Phi}^*$ in (33).[22] An implication of Proposition 1 is that liability for wrong treatments is essential for the expert to exert diagnostic effort. If $D_o = D_u = 0$, then $\overline{\Phi}^* = \overline{\Phi}_D = 0$, with the equilibrium markups being $\Phi_M^* = \Phi_m^* = 0$. This in turn implies that $\hat{k}(D,\Phi^*) = \overline{\Phi}^* = 0$, and the expert never invests in diagnosis even if k is small. However, a positive liability may also cause inefficiency, because it can discourage the expert from providing a treatment when he does not initially learn t and k is high.

From Lemma 4, when $\Delta C > \Delta C^*$, if $D_u = 0$, then $\Phi_m^* = (1-\theta)\alpha_u D_u$ and the expert would choose (N, T_m) instead of R if $k > \hat{k}(D, \Phi)$. On the other hand, when $\Delta C \le \Delta C^*$, if $D_o = 0$, then $\Phi_M^* = \theta\alpha_o D_o$, with $S_{NT} = S_{NM} > S_{Nm}$. This suggests that for different liabilities that have the same $\overline{\Phi}_D = \theta(1-\theta)(\alpha_o D_o + \alpha_u D_u)$, by setting either $D_u = 0$ or $D_o = 0$, we may induce the expert to avoid choices that fail to maximize S_{NT}. This observation is useful when we analyze optimal liabilities in the next section.

3.4 Numerical Example

We conclude this section by illustrating the market equilibrium for a given D with a numerical example. Assume $\theta = 0.5$, $z_o = 30$, $z_u = 120$, $\alpha_o = 0.2$, $\alpha_u = 0.9$, $C_M = 30$, $C_m = 10$, $x = 200$, and k is uniformly distributed on $[0, 200]$. Then $\overline{C} = \Delta C = 20 < \Delta C^* = (1-\theta)z_u - \theta z_o = 45$, $k^* = 25$, and the parameter values satisfy our assumption (2). At the efficient outcome, $\hat{k}(D, \Phi) = k^*$ and the expert chooses (N, T_M) if he does not initially learn t and $k > k^*$.

First, suppose $D_o = 320$ and $D_u = 40$, with $\theta\alpha_o D_o = 32 > (1-\theta)\,\alpha_u D_u = 18$ and $\overline{\Phi}_D = 25$.

[22] Although equation (33) is complicated, the equilibrium markups under a welfare-maximizing liability—as we shall see later—will be very simple: $\Phi_M^* = \theta\alpha_o D_o$ and $\Phi_m^* = (1-\theta)\alpha_u D_u$, with $\overline{\Phi}^* = \overline{\Phi}_D$.

Then

$$S\left(\overline{\Phi}\right)=\left[\beta+(1-\beta)\frac{\overline{\Phi}}{200}\right]\left(-20-\overline{\Phi}\right)+(1-\beta)\left[1-\frac{\overline{\Phi}}{200}\right]S_{NT}(\overline{\Phi}), \tag{22}$$

where

$$S_{NT}(\overline{\Phi})=\begin{cases} S_{NM}=-45 & if \quad \overline{\Phi}\in[16,25] \\ S_{Nm}=-70 & if \quad \overline{\Phi}\in[9,16) \\ -x=-200 & if \quad \overline{\Phi}\in[0,9) \end{cases}. \tag{23}$$

Figure 1 below illustrates $S(\overline{\Phi})$ for different values of β. When $\beta=0.2$ (left panel), $S\left(\overline{\Phi}\right)$ is maximized at $\overline{\Phi}^{*}=16$ with $(\Phi_M^*,\Phi_m^*)=(32,0)$, $\hat{k}(D,\Phi^*)=16<k^*$, and $S\left(\overline{\Phi}^*\right)=-42.6$; if the expert does not initially learn t and $k>\hat{k}(D,\Phi^*)$, he chooses R if $\overline{\Phi}\in[0,9)$, chooses (N,T_m) if $\overline{\Phi}\in[9,16)$, and chooses (N,T_M) if $\overline{\Phi}\in[16,25]$. When $\beta=0.8$ (right panel), $S\left(\overline{\Phi}\right)$ is maximized at $\overline{\Phi}^{*}=9$ with $(\Phi_M^*,\Phi_m^*)=(0,18)$, $\hat{k}(D,\Phi^*)=9$, and $S\left(\overline{\Phi}^*\right)=-36.8$; if the expert does not initially learn t and $k>\hat{k}(D,\Phi^*)$, his choices are the same as those under $\beta=0.2$.

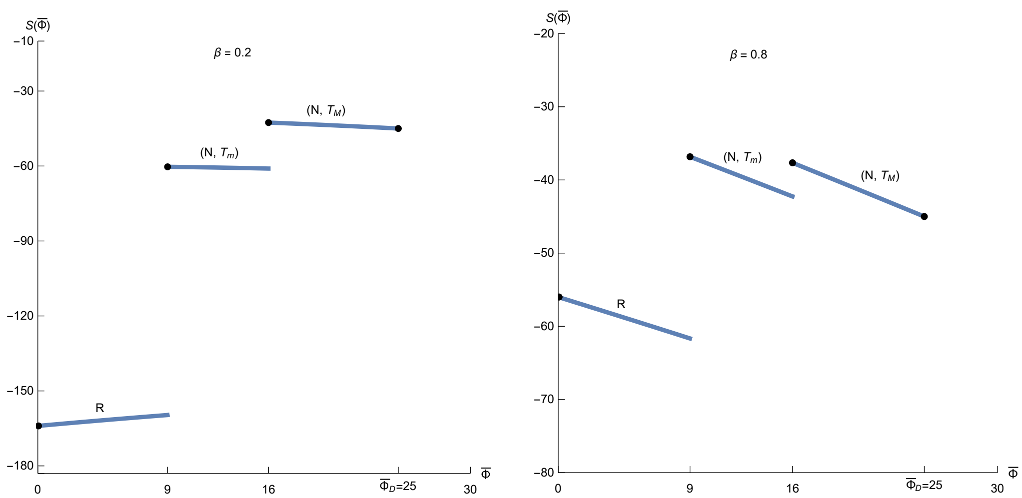

Figure 1: Consumer surplus $S(\overline{\Phi})$ given $(D_o,D_u)=(320,40)$.

Next, suppose $D_o=500$ and $D_u=0$, with $\theta\alpha_oD_o=50>(1-\theta)\,\alpha_uD_u=0$ and $\overline{\Phi}_D=25$.

Consumer surplus $S\left(\overline{\Phi}\right)$ is again given by (22) but now

$$S_{NT}(\overline{\Phi}) = \begin{cases} S_{NM} = -45 & if \quad \overline{\Phi} = 25 \\ S_{Nm} = -70 & if \quad \overline{\Phi} \in [0, 25) \end{cases}. \tag{24}$$

In Figure 2, when $\beta = 0.2$ (left panel), $S\left(\overline{\Phi}\right)$ is maximized at $\overline{\Phi}^* = \overline{\Phi}_D = 25$ with $\Phi_M^* = 50$ and $\Phi_m^* = 0$, $\hat{k}(D,\Phi^*) = 25 = k^*$, and $S\left(\overline{\Phi}^*\right) = -45$; if the expert does not initially learn t and $k > \hat{k}(D, \Phi^*)$, he chooses (N, T_m) if $\overline{\Phi} \in [0, 25)$ and chooses (N, T_M) if $\overline{\Phi} = 25$. When $\beta = 0.8$ (right panel), $S\left(\overline{\Phi}\right)$ is maximized at $\overline{\Phi}^* = 0$ with $(\Phi_M^*, \Phi_m^*) = (0, 0)$, $\hat{k}(D,\ \Phi^*) = 0$, and $S\left(\overline{\Phi}^*\right) = -30$; if the expert does not initially learn t and $k > \hat{k}(D, \Phi^*)$, his choices are the same as those under $\beta = 0.2$.

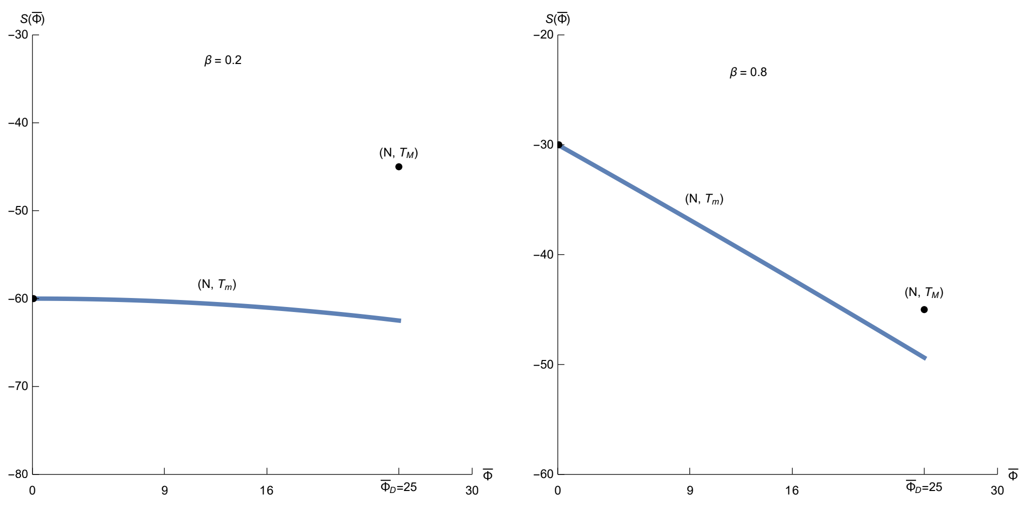

Figure 2: Consumer surplus $S(\overline{\Phi})$ given $(D_o, D_u) = (500, 0)$.

This example shows that under different liabilities, the equilibrium expected markup $(\overline{\Phi}^*)$, and its associated markups (Φ_M^*, Φ_m^*), may fall in different intervals in (23) and (24), which leads to different choices by the expert when he does not know t. The example also illustrates that both consumer surplus and welfare may vary with liability. In particular, when $\beta = 0.2$, $(D_o, D_u) = (500, 0)$ leads to the efficient outcome that maximizes welfare, whereas $(D_o, D_u) = (320, 40)$ fails to do so; but consumer surplus is lower under $(D_o, D_u) = (500, 0)$ than under $(D_o, D_u) = (320, 40)$. When $\beta = 0.8$, neither $(D_o, D_u) = (320, 40)$ nor $(500, 0)$ leads to the efficient outcome. Hence, not only equilibrium prices that maximize consumer surplus may not result in the efficient outcome, but also there is potentially a divergence between liabilities that maximize consumer surplus and those that maximize welfare.

4. Optimal Liability

This section analyzes optimal liabilities that maximize welfare, as measured by the sum of expected consumer surplus and profit (i.e., the expected total surplus), and provides conditions under which full efficiency may or may not be attained. In general, despite our maintained conditions (6) and (16), there is still a very large set of liabilities to be considered. The following result greatly simplifies our analysis by narrowing our search for optimal liabilities to a much smaller set.

Lemma 5 *In searching for an optimal liability rule, it suffices to consider any* $D = (D_o, D_u)$ *that induces equilibrium expected markup* $\overline{\Phi}^* = \overline{\Phi}_D$ *with*

$$\Phi_M^* = \theta\alpha_o D_o \quad \textit{and} \quad \Phi_m^* = (1-\theta)\alpha_u D_u. \tag{25}$$

Lemma 5 is proved based on the idea that if under some D condition (25) does not hold, then there is another liability rule under which (25) holds and also results in weakly higher welfare. Because we are interested in liability rules that maximize welfare, from now on we focus on liabilities under which (Φ_M^*, Φ_m^*) satisfy (25). That is, to find an optimal liability, we can focus on D that would induce $\overline{\Phi}^* = \overline{\Phi}_D$. Given any such D, under the equilibrium $\overline{\Phi}^*$ and with $\hat{k}(D, \Phi^*) = \overline{\Phi}^*$, welfare is

$$\begin{aligned} W(\hat{k}) = {} & (1-\beta)\left[1 - F(\hat{k})\right] \max\{-\theta z_o - C_M, \; -(1-\theta)z_u - C_m\} \\ & + \left[\beta + (1-\beta)F(\hat{k})\right] [-\theta C_m - (1-\theta)C_M] - (1-\beta)\int_0^{\hat{k}} kf(k)dk, \end{aligned} \tag{26}$$

where on the right-hand side of (26) the first term is the expected total surplus when the expert does not initially learn t and will also not incur k,[23] the second term is the expected total surplus when the expert knows t either initially or after incurring k, and the third term is the expected diagnostic cost. When D changes, $\overline{\Phi}^*$ and equilibrium welfare may both also change. An optimal liability D maximizes $W(\hat{k})$, and an efficient liability is an optimal liability that implements the first best described in Lemma 1. When an optimal liability fails to achieve the first best, we call it a second-best liability.

We next establish a necessary and sufficient condition for any efficient liability, and provide explicit parameter values under which full efficiency may or may not be attained under a welfare-maximizing liability. Recall that, from (18) and (5), $\overline{\Phi}_D = \theta(1-\theta)(\alpha_o D_o + \alpha_u D_u)$ and $k^* = \min\{\theta(\Delta C + z_o),\ (1-\theta)(z_u - \Delta C)\}$.

Proposition 2 *(1) The model has an equilibrium that achieves full efficiency if and only if there exists some $D^* = (D_o^*, D_u^*)$ such that* [24]

$$\text{(27)} \qquad (i)\ \overline{\Phi}^* = \overline{\Phi}_{D^*} \quad \text{and} \quad (ii)\ \overline{\Phi}_{D^*} = k^*.$$

(2) If D^* is an efficient liability with $D_o^* > 0$ and $D_u^* > 0$, then so is $D^* = \left(\frac{k^*}{\theta(1-\theta)\alpha_o}, 0\right)$ or $D^* = \left(0, \frac{k^*}{\theta(1-\theta)\alpha_u}\right)$.

(3) $D^* = \left(\frac{\Delta C + z_o}{(1-\theta)\alpha_o}, 0\right)$ satisfies (27) if $\Delta C \leq \Delta C^*$ and z_u is sufficiently large, while $D^* = \left(0, \frac{z_u - \Delta C}{\theta\alpha_u}\right)$ satisfies (27) if $\Delta C > \Delta C^*$ and z_o is sufficiently large; however, there is $\hat{\beta} < 1$ such that if $\beta > \hat{\beta}$, no efficient liability exists.

Proposition 2 states that the existence of a liability satisfying $\overline{\Phi}_D = k^*$ is a necessary condition for full efficiency in equilibrium. For such a liability, if its induced expected equilibrium markup satisfies $\overline{\Phi}^* = \overline{\Phi}_D = k^*$, then full efficiency is achieved. If instead $\overline{\Phi}^* < \overline{\Phi}_D$, then the equilibrium is not fully efficient. This latter case can arise if the consumer finds it optimal to set lower prices in order to reduce the expert's information rents so that for any D satisfying $\overline{\Phi}^* = \overline{\Phi}_D$, $\overline{\Phi}_D < k^*$. Proposition 2 further describes explicitly two specific liabilities that are efficient under certain conditions and also a situation where an efficient liability fails to exist. In particular, if the

[23] In this case, the equilibrium markups will lead to $\max\{S_{NM}, S_{Nm}\}$ for the consumer. When (25) holds, $W(N, T_M) = S_{NM} = -\theta z_o - C_M$ and $W(N, T_m) = S_{Nm} = -(1-\theta)z_u - C_m$.

[24] Condition (27) can be alternatively written as $\overline{\Phi}_{D^*} = k^* \in \arg\max_{\overline{\Phi} \in [0, \overline{\Phi}_{D^*}]} S(\overline{\Phi})$.

loss from a wrong treatment is sufficiently high, either $D^* = \left(\frac{\Delta C + z_o}{(1-\theta)\alpha_o}, 0\right)$ or $D^* = \left(0, \frac{z_u - \Delta C}{\theta \alpha_u}\right)$ will satisfy (27) in equilibrium, whereas if β is sufficiently close to 1, no efficient liability exists. Moreover, Proposition 2 implies that to find an efficient liability, we can focus on these two liabilities because one of them is efficient whenever an efficient liability exists.

The efficient liabilities in Proposition 2, $D^* = \left(\frac{\Delta C + z_o}{(1-\theta)\alpha_o}, 0\right)$ or $D^* = \left(0, \frac{z_u - \Delta C}{\theta \alpha_u}\right)$, suggest that D_o^* or D_u^* is higher when the loss from a wrong treatment (z_o or z_u) is higher or the probability that a loss is verified (α_o or α_u) is lower. To illustrate why $D^* = \left(\frac{\Delta C + z_o}{(1-\theta)\alpha_o}, 0\right)$ is an efficient liability when $\Delta C \leq \Delta C^*$ and z_u is high, notice that under D^*, $\overline{\Phi}_{D^*} = k^*$ and D^* is efficient if it induces $\overline{\Phi}^* = \overline{\Phi}_{D^*}$. Since $D_u^* = 0$, $\Phi_m^* = (1-\theta)\alpha_u D_u^*$ always holds.[25] Thus, if $\overline{\Phi}^* < \overline{\Phi}_{D^*}$, it must be $\Phi_M^* < \theta \alpha_o D_o^*$; then, when the expert does not know t, he will choose (N, T_m) and the consumer will receive $-C_m - (1-\theta) z_u$, which decreases if z_u increases. Therefore, if z_u is high enough, it would be optimal for the consumer to set $\Phi_M^* = \theta \alpha_o D_o^*$ with $\overline{\Phi}^* = \overline{\Phi}_{D^*}$ to avoid the loss from undertreatment, and hence D^* induces $\overline{\Phi}^* = \overline{\Phi}_{D^*} = k^*$. Similarly, when $\Delta C > \Delta C^*$ and z_o is high enough, under $D^* = \left(0, \frac{k^*}{\theta(1-\theta)\alpha_u}\right)$ the consumer will optimally set $\Phi_m^* = (1-\theta)\alpha_u D_u^*$ to avoid overtreatment, and hence D^* also induces $\overline{\Phi}^* = \overline{\Phi}_{D^*} = k^*$.

To see the intuition behind the failure of full efficiency when $\beta > \hat{\beta}$, notice that in choosing the optimal prices, the consumer faces the trade-off between extracting more surplus when the expert knows t by reducing the prices, and incentivizing the expert to (i) exert diagnostic effort or (ii) provide service even if he does not learn t when k is high. In order to reduce the expert's information rent, the consumer may want to reduce the prices below the level that would induce the efficient effort. This problem is exacerbated when the consumer can be (partially) compensated by liability for the loss associated with an inappropriate treatment and hence she does not bear the full social cost of the loss. When β is high so that the expert will choose the right treatment sufficiently often even without incurring k (and with D_u or $D_o = 0$ he will not choose R), the consumer will indeed optimally choose inefficiently low prices that result in deficient diagnostic effort. As a result, when β is sufficiently high, an efficient liability—one that

[25] When $\Delta C \leq \Delta C^*$, the consumer would like the expert to choose (N, T_M) if he must choose among $\{(N, T_M), (N, T_m)\}$. Thus, if $D_u > 0$, the consumer has more incentive to deviate to $\Phi_m < (1-\theta)\alpha_u D_u^*$ under which the expert will not choose (N, T_m). Setting $D_u^* = 0$ prevents the deviation so that part (i) of (27) is easier to satisfy, even though there are numerious D with $\overline{\Phi}_D = k^*$.

will ensure $\hat{k}(D, \Phi^*) = \overline{\Phi}^* = k^*$—fails to exist.

Proposition 2 highlights the subtlety in the design of an optimal liability: while liability is necessary to provide incentives for the expert to exert effort, it also creates a divergence between the social and private costs of a loss to the consumer, as well as possibly causing the expert to provide a wrong treatment or decline to serve the consumer. Although there can be more than one D such that $\overline{\Phi}_D = k^*$, it requires great care to find a D^* which would also induce the consumer to set $\overline{\Phi}^* = \overline{\Phi}_{D^*}$ while ensuring $\overline{\Phi}_{D^*} = k^*$, and such a D^* does not always exist. Therefore, while unconstrained competition in the expert market maximizes consumer surplus, inefficiency may arise even under an optimal liability.

When an efficient liability—the first-best liability—does not exist, we can still find a second-best liability that maximizes welfare in equilibrium.

Corollary 1 *When the first-best liability fails to exist, there is a second-best liability* $D^{**} = (D_o^{**}, D_u^{**})$ *that maximizes welfare. The equilibrium expected markup under* D^{**} *satisfies* $\overline{\Phi}^* = \overline{\Phi}_{D^{**}} = \theta(1-\theta)(\alpha_o D_o^{**} + \alpha_u D_u^{**})$ *with* $\Phi_M^* = \theta\alpha_o D_o^{**}$ *and* $\Phi_m^* = (1-\theta)\,\alpha_u D_u^{**}$*, but the expert's diagnostic effort is below the efficient level:* $\hat{k}(D^{**}, \Phi^*) = \overline{\Phi}_{D^{**}} < k^*$.

We can briefly illustrate Proposition 2 and Corollary 1 with the example in Section 3.4, in which $z_u = 120$. If $\beta = 0.2$, $D = \left(\frac{k^*}{\theta(1-\theta)\alpha_o}, 0\right) = (500, 0)$ induces $\overline{\Phi}^* = \overline{\Phi}_D = k^* = 25$ with full efficiency. But if $\beta = 0.8$, $D = (500, 0)$ induces $\overline{\Phi}^* = 0 < \bar{\Phi}_D$ with $\Phi_M^* < \theta\alpha_o D_o$ and the equilibrium is not fully efficient; in fact, this is a case where there exists no efficient liability.[26] The second-best liability is $(D_o^{**}, D_u^{**}) = (127.98, 0)$, which leads to the highest expected markup that satisfies (25): $\overline{\Phi}^* = \overline{\Phi}_{D^{**}} = 6.399$ with $\Phi_M^* = \theta\alpha_o D_o^{**} = 12.798$ and $\Phi_m^* = 0$, under which the expert chooses (N, T_M) when he does not know t. Then, the consumer has no incentive to lower Φ_M below Φ_M^*, because $\Phi_M < \theta\alpha_o D_o^{**}$ would cause the expert to choose (N, T_m) instead of (N, T_M), which would not increase consumer surplus. Since $\hat{k}(D^{**}, \Phi^*) = \overline{\Phi}^* = 6.399 < k^*$, the diagnostic effort is lower than the efficient level.

Therefore, when an optimal liability is unable to achieve full efficiency, it can implement the second-best outcome by inducing the highest possible equilibrium prices under which the expert

[26] But if z_u is much larger, for example $z_u = 320$, then full efficiency can again be achieved.

expects the same (zero) profit from both treatments when he does not learn t; the expert will then choose the treatments efficiently in such situations, but the prices are not high enough to motivate the expert to exert efficient diagnostic effort. In particular, if β is sufficiently high, or in the case of pure credence goods with $\alpha_o \to 0$ and $\alpha_u \to 0$, given liabilities satisfying (16), there exist no equilibrium prices such that condition (27) holds. Then, full efficiency cannot be achieved and the expert chooses insufficient diagnostic effort under a second-best liability.

Notice that an optimal liability $\hat{D} = \left(\hat{D}_o, \hat{D}_u\right)$—be it the first best or second best—can be expressed as a multiplier of the loss from undertreatment or overtreatment: $\hat{D}_o = \gamma_o z_o$ and $\hat{D}_u = \gamma_u z_u$. For instance, for the efficient liabilities $D^* = \left(\frac{k^*}{\theta(1-\theta)\alpha_o}, 0\right)$ or $D^* = \left(0, \frac{k^*}{\theta(1-\theta)\alpha_u}\right)$ in Proposition 2, the multipliers are respectively

$$\gamma_o = \frac{k^*}{\theta(1-\theta)\alpha_o z_o} \text{ and } \gamma_u = 0; \quad \text{or} \quad \gamma_o = 0 \text{ and } \gamma_u = \frac{k^*}{\theta(1-\theta)\alpha_u z_u}.$$

It is possible that $\gamma_u > 1$ or $\gamma_o > 1$; that is, there might be punitive damages.[27]

Moreover, we may consider liability policies in which the expert is liable to pay only a fixed portion γ of the verified loss, with $D = \gamma z_i$ for $i = o, u$. Such a policy can achieve the same welfare as an optimal liability $\hat{D}$ with equilibrium markups $\left(\hat{\Phi}_M^*, \hat{\Phi}_m^*\right)$ if γ induces $\Phi_M^* = \theta\alpha_o\gamma z_o = \hat{\Phi}_M^*$ and $\Phi_m^* = (1-\theta)\,\alpha_u\gamma z_u = \hat{\Phi}_m^*$ with equilibrium

$$\overline{\Phi}^* = \theta\,(1-\theta)\,(\alpha_o\gamma z_o + \alpha_u\gamma z_u) = \hat{k}\left(\hat{D}, \hat{\Phi}^*\right).$$

However, it appears that restricting to a common γ, under which $\Phi_M^* > 0$ and $\Phi_m^* > 0$, will significantly limit the ability to maximize welfare. In particular, the two efficient liabilities in Proposition 2 contain either $D_u^* = 0$ or $D_o^* = 0$, under which $\Phi_m^* = 0$ or $\Phi_M^* = 0$, to prevent the consumer from deviating to a lower Φ_m or Φ_M. This is not possible under any $\gamma > 0$.

[27] Although punitive damages are rare, they are sometimes awarded in practice. For example, in its opinion filed on March 2, 2021, the Missouri Supreme Court affirmed a Missouri medical malpractice jury's award in favor of the plaintiff in the amounts of \$269,780.80 for economic damages, \$300,000 for noneconomic damages, and \$300,000 for aggravating circumstances damages (punitive damages). See https://medicalmalpracticelawyers.com/missouri-supreme-court-affirms-punitive-damages-award-in-medical-malpractice-case/.

5. Remedies to Restore Efficiency

As we demonstrated in Section 4, equilibrium in the expert market need not be efficient, even when liability is optimally chosen. We now show that there are two potential remedies that can restore full efficiency: regulating prices or imposing the obligation for the expert to serve.

5.1 Minimum Price Constraint

Suppose there is a minimum-price regulation that, for given $D = (D_o, D_u)$, requires[28]

$$\Phi_M \geq \theta\alpha_o D_o, \quad \Phi_m \geq (1-\theta)\alpha_u D_u. \tag{28}$$

When (28) is satisfied, Φ_M and Φ_m are high enough so that the expert, whose outside option is zero profit, will receive nonnegative expected profits from providing each treatment even without knowing t, i.e., $\pi(N, T_M) \geq 0$ and $\pi(N, T_m) \geq 0$. In equilibrium, constraint (28) and Lemma 3 imply

$$\Phi_M^* = \theta\alpha_o D_o, \quad \Phi_m^* = (1-\theta)\alpha_u D_u. \tag{29}$$

The result below establishes that there exists an optimal liability rule that induces the efficient outcome in equilibrium.

Proposition 3 *Suppose that (28) holds. Then, the following liability rule results in the efficient outcome in equilibrium*

$$D_u^* = \frac{k^*}{(1-\theta)\alpha_u}, \qquad D_o^* = \frac{k^*}{\theta\alpha_o}. \tag{30}$$

Notice that with the liability rule that implements the efficient outcome, the ex ante expected equilibrium markup $\overline{\Phi}^*$ is equal to the efficient critical value k^* for diagnostic effort. Thus, while

[28] This minimum-price constraint may also arise without regulation if the expert and the consumer share bargaining power in setting prices because, for example, the consumer has costs to compare prices from potential service providers. The expert may then be able to insist on prices that would ensure nonnegative profits for each treatment as in (28).

there exist a range of liabilities that would induce the markups given in (29) for the two treatments and these markups generally differ, they all have the same expected value under the efficient liability, being equal to the expert's expected liability cost from treating the consumer without knowing t. By selecting D^* that equates this expected liability cost to the efficient k^*, the efficient liability incentivizes the expert to fully internalize the social benefit from choosing the efficient diagnostic effort. In our main model without the restriction on prices, the consumer may not fully internalize this benefit and may thus set too low prices for the expert, inefficiently reducing his effort. The minimum-price constraint removes this possibility and helps restore efficiency.[29]

5.2 Obligation to Serve

We now show that, instead of price regulation, full efficiency can also be restored with a properly-chosen liability rule if there is regulation on the expert's obligation to serve. Specifically, suppose that upon seeing the consumer, the expert is not allowed to choose R. Then, if the expert agrees to see the consumer, he will incur k if and only if

$$\pi(E_T) = \theta\Phi_m + (1-\theta)\Phi_M - k \geq \max\{\pi(N, T_M), \pi(N, T_m)\},$$

or, equivalently,

$$k \leq \hat{k}(D, \Phi) = \min\{\theta(\Phi_m - \Phi_M + \alpha_o D_o), (1-\theta)(\Phi_M - \Phi_m + \alpha_u D_u)\}.$$

Given (D_o, D_u), the consumer chooses (Φ_M, Φ_m) to maximize her expected surplus, subject to the constraints that the expert will incur k if and only if $k \leq \hat{k}(D, \Phi)$ and that he receives nonnegative expected profit by agreeing to see the consumer.

The expert is willing to accept (Φ_M, Φ_m) with the obligation to serve if his expected profit is

[29] To use the minimum-price regulation, a regulator would need to know the markups, which could be difficult in practice. When the regulator only knows that the markups belong to some ranges, applying the restriction to the highest markups in the ranges could ensure efficiency, though this would reduce consumer surplus. We discuss this potential conflict between efficiency and consumer surplus when considering alternative pricing regimes in Section 6.

nonnegative:

$$\Pi(D,\Phi) = \left[\beta + (1-\beta)F(\hat{k})\right][\theta\Phi_m + (1-\theta)\Phi_M] - (1-\beta)\int_0^{\hat{k}} tdF(t)$$
$$+ (1-\beta)\left[1 - F(\hat{k})\right]\max\{\Phi_M - \theta\alpha_o D_o, \Phi_m - (1-\theta)\alpha_u D_u\} \geq 0,$$

where the first term is the expert's expected payoff when he knows t and provides the right treatment, the second term is his expected diagnosis cost anticipating that he will exert diagnostic effort if and only if $k \leq \hat{k}$, and the last term is the expert's expected payoff when he does not know t and makes a choice between (N, T_M) and (N, T_m). The result below identifies a liability rule that achieves full efficiency in equilibrium when the expert is obligated to serve the consumer.

Proposition 4 *Suppose that the expert is obligated to treat the consumer after seeing her. Then, liability rule D^* with*

(31) $$D_o^* = \frac{k^*}{\theta\alpha_o}, \qquad D_u^* = \frac{k^*}{(1-\theta)\,\alpha_u}$$

induces the equilibrium markups

(32) $$\Phi_M^* = (1-\beta)\left[k^* - \int_0^{k^*} F(t)dt\right] = \Phi_m^*$$

and achieves the efficient outcome.

Under (D_o^*, D_u^*) in (31), the markups (Φ_M^*, Φ_m^*) in (32) satisfy

$$\Phi_M^* - \theta\alpha_o D_o^* = \Phi_m^* - (1-\theta)\,\alpha_u D_u^*$$

and hence $\pi(N, T_M) = \pi(N, T_m)$. The expert will thus choose between (N, T_M) and (N, T_m) that favors the consumer when he does not know t. Moreover, under these markups $\hat{k}(D^*, \Phi^*) = \theta(1-\theta)(\alpha_o D_o^* + \alpha_u D_u^*) = k^*$, and thus the expert's diagnostic effort is also efficient. The markups in (32) drive the expert's expected profit to zero and thus must be optimal for the consumer. Therefore, D^* induces the efficient outcome in equilibrium.

When the expert must provide a treatment upon seeing the consumer even if he has no precise information about the consumer's type, the obligation to serve restores efficiency by eliminating the inefficiency that arises when the prices are not high enough to motivate the expert to treat the consumer if he does not learn t. Essentially, this requirement enables the consumer to extract all information rents from the expert. In reality, the obligation to serve may be imposed under certain situations, such as for emergency care. However, in other cases, it may be difficult to enforce the obligation to serve. After an initial consultation, it would seem reasonable that the expert, without taking any payment from the consumer, will have the right not to provide treatment. A dentist, for example, may simply refer a patient to a "specialist" after seeing her.

6. Discussion of Price Regimes

Our main model assumes that the consumer sets the prices for the expert's service. One possibility is that there is a large buyer organization that has the market power to set prices on behalf of individual consumers. In the healthcare industry, for instance, a big employer such as a university or a health maintenance organization may be in a position to do so.

As we pointed out earlier, an alternative interpretation of this price regime is that there is ex ante perfect price competition among multiple potential experts. For example, suppose that there are two ex ante homogeneous experts, A and B, who compete by simultaneously setting prices (P^i_M, P^i_m), with respective markups (Φ^i_M, Φ^i_m), $i \in \{A, B\}$. Then, from the familiar logic of Bertrand competition, the equilibrium prices must maximize the consumer's surplus. Thus, to find welfare-maximizing liabilities, we can again focus on any liability $D = (D_o, D_u)$ such that for $i = A, B$, $(\Phi^i_M, \Phi^i_m) = (\theta\alpha_o D_o, (1-\theta)\alpha_u D_u)$, with the expected markup $\overline{\Phi}^i = \theta(1-\theta)(\alpha_o D_o + \alpha_u D_u)$. If D^* is an efficient liability in equilibrium when the consumer sets prices, then D^* is also an efficient liability in the equilibrium under this alternative assumption of price formation through ex ante perfect competition. Similarly, for the case where the first best cannot be achieved, there exists some D^{**} that is a second-best liability both under price-setting by the consumer and under ex ante perfect price competition among experts.[30]

[30] One potential complication in the presence of multiple experts is that, ex post, if an expert refuses to treat the consumer after seeing her, the consumer may switch to another expert. We implicitly assume that once a consumer

Is the assumption that the consumer sets prices crucial for our main results on the role of liability in expert markets? To gain insight on this issue, we next consider the other polar case in which a monopolist expert sets prices (P_M, P_m)—or equivalently markups (Φ_M, Φ_m)—respectively for treatments T_M and T_m, before the consumer's visit. The expert now chooses the prices to maximize his expected profit. As we show in Proposition 5 below, a liability rule $D^* = (D_o^*, D_u^*)$ that satisfies $\overline{\Phi}_{D^*} = \theta(1-\theta)(\alpha_o D_o^* + \alpha_u D_u^*) = k^*$, which is only part (ii) of condition (27), will now implement the efficient outcome. Therefore an efficient liability always exists when the expert sets prices.

The basic idea to prove Proposition 5 is as follows: Denote the first-best welfare, as described in Lemma 1, by W^*. Also notice that the consumer's reservation utility is $-x$, which is what she will obtain if not receiving a treatment from the monopoly expert. If we can demonstrate that, under some liability rule and the associated monopoly prices, W^* is achieved while the consumer receives $-x$, then the outcome must be both efficient and profit-maximizing for the expert. Such a liability rule then implements the efficient outcome.

Proposition 5 *Suppose that a monopoly expert sets prices. Then, a liability rule* $D^* = (D_o^*, D_u^*)$ *that satisfies* $\theta(1-\theta)(\alpha_o D_o^* + \alpha_u D_u^*) = k^*$ *implements the efficient outcome in equilibrium.*

Proposition 5 suggests that in expert markets, welfare can be higher when prices are set by a monopoly seller rather than by competitive sellers. While this may appear intriguing at first glance, it actually has a simple and intuitive explanation: because the expert is the party who has the private information and may need to exert costly efforts, giving the pricing power to the expert enables him to fully appropriate the efficiency gains from his actions and, consequently, motivates him to make the efficient effort choice.

However, under both price regimes, a well-designed liability rule is essential for market efficiency. In particular, if $D_o = D_u = 0$ (i.e., without liability), the market is generally not efficient. Moreover, the relations between the markups bear the same pattern under the two price regimes. Furthermore, the optimal liabilities under both price regimes generally vary with the environment

accepts an expert's prices and is seen/examined by him, she cannot gain by seeing another expert, possibly because the experts have the same need for extra diagnostic efforts to discover t and identical values of k.

in predictable ways. For instance, when the loss from a wrong treatment is higher, the optimal liabilities are also higher to increase the incentive for diagnostic efforts.

Although under both price regimes a well-designed liability rule can improve efficiency and an efficient liability always exists when the expert sets prices, consumer surplus is much higher when the consumer sets prices. Therefore, potentially there is a conflict between price regimes that are best for consumers and those that are best for efficiency. As we analyzed earlier in Section 5, when prices are restricted to be above some minimum levels, there is always an efficient liability that implements the first-best outcome, while consumer surplus is maximized subject to this minimum-price constraint. If we interpret such a price regime where prices are neither perfectly competitive nor monopolistic as arising from some (small) market frictions, then our results suggest that such frictions can improve market efficiency while providing the highest possible consumer surplus under the efficiency constraint. In other words, if we value both consumer surplus and efficiency, an expert market with some frictions in price competition can outperform one in which there is perfect price competition among (potential) experts.[31]

7. Conclusion

This paper has studied efficient liability in a model of expert markets where proper incentives are needed for the expert to exert diagnostic efforts and to recommend the appropriate treatment. We characterize the necessary and sufficient condition under which there exists a liability rule to implement full efficiency in equilibrium, and identify situations where the condition is satisfied or may fail. An efficient liability rule imposes a penalty on the expert that increases with verified consumer loss from overtreatment or undertreatment but decreases with the probability that a loss can be verified. We also show that while liability is necessary to provide incentives to the expert, it creates a divergence between the social and private costs of a loss to the consumer. Consequently, unfettered price competition between experts, while maximizing consumer surplus, can render it impossible to achieve full efficiency. Under a second-best liability rule, the expert

[31] This point has an interesting connection to the results in other market settings. For instance, in markets with consumer search and product quality differences, both consumer surplus and efficiency can achieve their highest levels when there are some (small) entry costs by sellers (Chen and Zhang, 2018) or—for experience goods—some (small) search costs by consumers (Chen et al., in press).

generally under-invests in diagnostic efforts. A (regulatory) constraint on minimum prices or on the obligation to serve enables an optimal liability rule to restore full efficiency. An efficient liability always exists also when a monopoly expert sets the treatment prices, though in this case consumer surplus would be minimized.

We have analyzed a stylized model. There are other factors that can potentially impact the performance of expert markets. For example, if there are repeat purchases, reputation concerns can motivate experts to exert effort and behave honestly in serving consumers. But reputation may be fragile, and a well-designed liability rule can achieve efficiency even when reputation does not. It is also possible that the expert and the consumer will rely on private contracts instead of legal liability for damage payments in the case of a consumer loss; in such situations, we may interpret the optimal liability in our model as privately stipulated damages. However, private contracting for damage payments can have high transaction costs and contract enforcement may still rely on the legal system. Moreover, if the damage payments through private contracting are designed by the consumer or offered by (perfectly) competitive experts, they will generally differ from the welfare-maximizing liability because of the difference between social and private costs from maltreatment.

In some situations, the compensation to the consumer for the loss from a wrong treatment may differ from the expert's liability. Recall that in our baseline setting, the source of inefficiency is the consumer's incentive to choose inefficiently low prices to extract information rents from the expert, and with liability this incentive problem is exacerbated because the consumer then does not bear the full cost of the loss. Therefore, if the liability payments do not directly go to the consumer, the consumer will have less incentive to set inefficiently low prices. This could improve efficiency, but may still not restore full efficiency because the consumer remains to have an incentive to lower prices in order to reduce the expert's information rents. A related issue is liability insurance, under which the expert does not bear the full cost of liability. This would undermine the role of liability in motivating the expert to exert diagnostic effort, thereby reducing efficiency; but it could also increase the incentive for the expert to serve consumers and, in a more general model with risk averse experts, permit efficient risk sharing. These and other policy issues concerning the design of liability in expert markets deserve more analysis in future research.

The difficulties in providing proper incentives to experts (such as physicians and dentists) are well known. The fact that malpractice liabilities are a prominent feature of markets such as those for healthcare further suggests that legal liability plays an important role in expert markets.[32] By showing how an efficient liability can be designed in a model of adverse selection and moral hazard, this paper offers new insights on improving the performance of expert markets.

Appendix

The appendix contains proofs for Lemmas 4-5, Propositions 2-5 and Corollary 1.

Proof of Lemma 4. First, we derive the consumer's optimal markups $(\hat{\Phi}_M, \hat{\Phi}_m)$ as a function of $\overline{\Phi}$ in the following:

(33)

$$\begin{array}{ll} \hat{\Phi}_M = \theta\alpha_o D_o & \text{if } \frac{\overline{\Phi}}{\theta(1-\theta)} \geq \alpha_o D_o \text{ and } \Delta C \leq \Delta C^*, \text{ or } \alpha_o D_o \leq \frac{\overline{\Phi}}{\theta(1-\theta)} < \alpha_u D_u, \text{ or } \overline{\Phi} = \overline{\Phi}_D \\ \hat{\Phi}_m = (1-\theta)\alpha_u D_u & \text{if } \frac{\overline{\Phi}}{\theta(1-\theta)} \geq \alpha_u D_u \text{ and } \Delta C > \Delta C^*, \text{ or } \alpha_u D_u \leq \frac{\overline{\Phi}}{\theta(1-\theta)} < \alpha_o D_o, \text{ or } \overline{\Phi} = \overline{\Phi}_D \\ \frac{\hat{\Phi}_M}{\theta} < \alpha_o D_o, \frac{\hat{\Phi}_m}{(1-\theta)} < \alpha_u D_u & \text{if } \frac{\overline{\Phi}}{\theta(1-\theta)} < \min\{\alpha_o D_o, \alpha_u D_u\} \end{array},$$

where $(1-\theta)\hat{\Phi}_M + \theta\hat{\Phi}_m = \overline{\Phi} \in [0, \overline{\Phi}_D]$. Recall that with $S_{NM} = -C_M - \theta z_o$ and $S_{Nm} = -C_m - (1-\theta)z_u$, part (i) of (2) implies $\min\{S_{NM}, S_{Nm}\} \geq -x$.

Notice that given $\overline{\Phi}$, the individual values of Φ_M and Φ_m matter only when the expert does not initially learn t and also chooses not to incur k. The optimal Φ_M and Φ_m are chosen to maximize the consumer surplus in this case, denoted by S_{NT}, as follows:

(i) If $\overline{\Phi} \geq \theta(1-\theta)\alpha_o D_o$ and $\Delta C \leq \Delta C^*$, then $S_{NM} \geq S_{Nm}$ and $\hat{\Phi}_M = \theta\alpha_o D_o$ maximizes S_{NT}. If $\theta(1-\theta)\alpha_o D_o \leq \overline{\Phi} < \theta(1-\theta)\alpha_u D_u$ or if $\overline{\Phi} = \overline{\Phi}_D$, then $\hat{\Phi}_M = \theta\alpha_o D_o$ also maximizes S_{NT} whether or not $\Delta C \leq \Delta C^*$. Notice that the only possible situations not covered above are either $\overline{\Phi} < \theta(1-\theta)\alpha_o D_o$ or $\theta(1-\theta)\max\{\alpha_o D_o, \alpha_u D_u\} \leq \overline{\Phi} < \overline{\Phi}_D$ and $\Delta C > \Delta C^*$.

(ii) If $\overline{\Phi} \geq \theta(1-\theta)\alpha_u D_u$ and $\Delta C > \Delta C^*$, then $S_{Nm} \geq S_{NM} \geq -x$ and $\hat{\Phi}_m = (1-\theta)\alpha_u D_u$ maximizes S_{NT}. If $\theta(1-\theta)\alpha_u D_u \leq \overline{\Phi} < \theta(1-\theta)\alpha_o D_o$ or if $\overline{\Phi} = \overline{\Phi}_D$, then $\hat{\Phi}_m = (1-\theta)\alpha_u D_u$ also maximizes S_{NT} even when $\Delta C \leq \Delta C^*$.

[32] In our model, the expert perfectly observes the consumer's problem with diagnosis effort. Our main results can still hold if we extend the model to a setting where the expert obtains only a noisy signal about the nature of the consumer's problem.

(iii) If $\overline{\Phi} < \theta(1-\theta)\min\{\alpha_o D_o, \alpha_u D_u\}$, then $\hat{\Phi}_M < \theta\alpha_o D_o$ and $\hat{\Phi}_m < (1-\theta)\alpha_u D_u$.

We have thus shown that for any given $\overline{\Phi}$, the optimal markups $(\hat{\Phi}_M, \hat{\Phi}_m)$ are given by (33).

Next, $S_{NT} = S_{NT}\left(\overline{\Phi}\right)$ follows directly from the derivation of equation (33). Notice that $\overline{\Phi} = \overline{\Phi}_D$ appears in (33) but not explicitly in (20), because $S_{NT} = S_{NM}$ if $\overline{\Phi} = \overline{\Phi}_D$ and $\Delta C \leq \Delta C^*$, while $S_{NT} = S_{Nm}$ if $\overline{\Phi} = \overline{\Phi}_D$ and $\Delta C > \Delta C^*$.

Finally, the expert will choose T_t for $t \in \{m, M\}$ with probability $\left[\beta + (1-\beta)F(\overline{\Phi})\right]$, and he does not initially learn t and also chooses not to incur k with probability $(1-\beta)\left[1 - F(\overline{\Phi})\right]$. Therefore, the consumer's expected surplus as a function of $\overline{\Phi}$, $S\left(\overline{\Phi}\right)$, is given by (19). ■

Proof of Lemma 5. Recall from Lemma 3 and Lemma 4 that $0 \leq \Phi_M^* \leq \theta\alpha_o D_o$ and $0 \leq \Phi_m^* \leq (1-\theta)\alpha_u D_u$. We show that if $\Phi_M^* \neq \theta\alpha_o D_o$ or $\Phi_m^* \neq (1-\theta)\alpha_u D_u$ under D, then there is another liability rule $\tilde{D}$, with $\tilde{\Phi}_M^* = \theta\alpha_o \tilde{D}_o$ and $\tilde{\Phi}_m^* = (1-\theta)\alpha_u \tilde{D}_u$, that leads to weakly higher welfare.

First, if $\Phi_M^* = \theta\alpha_o D_o$ but $\Phi_m^* < (1-\theta)\alpha_u D_u$ under D, then $S_{NT}(\overline{\Phi}) = S_{NM}$ by (20), $D_u > 0$, and $\overline{\Phi}^* < \overline{\Phi}_D$. Consider an alternative liability $\tilde{D} = (\tilde{D}_o, \tilde{D}_u)$ with $\tilde{D}_o = 0$ and $\tilde{D}_u = \frac{\overline{\Phi}^*}{\theta(1-\theta)\alpha_u}$. Under $\tilde{D}$, $\tilde{\Phi}_M^* = \theta\alpha_o \tilde{D}_o = 0$. Moreover, under $\tilde{D}$, if $\tilde{\Phi}_m < (1-\theta)\alpha_u \tilde{D}_u$, then $\tilde{S}_{NT}\left(\overline{\Phi}\right) = S_{NM}$ and $\tilde{S}(\overline{\Phi}) = S\left(\overline{\Phi}\right)$; while if $\tilde{\Phi}_m = (1-\theta)\alpha_u \tilde{D}_u = \frac{\overline{\Phi}^*}{\theta}$, then

$$\tilde{S}_{NT}\left(\overline{\Phi}\right) = \max\{S_{NM}, S_{Nm}\}$$

and $\tilde{S}\left(\overline{\Phi}\right) \geq S\left(\overline{\Phi}\right)$. Therefore $\tilde{\Phi}_m^* = (1-\theta)\alpha_u \tilde{D}_u$, and both consumer surplus and profit are weakly higher under $\tilde{D}$ than under D.

Next, if $\Phi_m^* = (1-\theta)\alpha_u D_u$ but $\Phi_M^* < \theta\alpha_o D_o$, then $S_{NT}\left(\overline{\Phi}\right) = S_{Nm}$, $D_o > 0$, and $\overline{\Phi}^* < \overline{\Phi}_D$. Consider an alternative liability $\tilde{D} = (\tilde{D}_o, \tilde{D}_u)$ with $\tilde{D}_o = \frac{\overline{\Phi}^*}{\theta(1-\theta)\alpha_o}$ and $\tilde{D}_u = 0$. Similarly as above, the consumer's optimal markups under $\tilde{D}$ will be $\tilde{\Phi}_M^* = \theta\alpha_o \tilde{D}_o = \frac{\overline{\Phi}^*}{1-\theta}$ and $\tilde{\Phi}_m^* = 0$, and welfare is weakly higher under $\tilde{D}$ than under D.

Finally, if $\Phi_M^* < \theta\alpha_o D_o$ and $\Phi_m^* < (1-\theta)\alpha_u D_u$, then $S_{NT}(\overline{\Phi}) = -x$. Consider an alternative liability $\tilde{D} = (\tilde{D}_o, \tilde{D}_u)$ with $\tilde{D}_o = \frac{\Phi_M^*}{\theta\alpha_o}$ and $\tilde{D}_u = D_u$. Then, the same (Φ_M^*, Φ_m^*) would still be optimal for the consumer, but now with $\Phi_M^* = \theta\alpha_o \tilde{D}_o$, $\tilde{S}_{NT}(\overline{\Phi}) = S_{NM} \geq -x = S_{NT}(\overline{\Phi})$, and thus welfare is weakly higher under $\tilde{D}$ than under D. Moreover, from the arguments above,

$\tilde{D}$ is still weakly dominated by another liability rule $\hat{D} = (\hat{D}_o, \hat{D}_u)$ under which $\overline{\Phi}^* = \overline{\Phi}_{\hat{D}}$ with $\Phi_M^* = \theta\alpha_o\hat{D}_o$ and $\Phi_m^* = (1-\theta)\alpha_u\hat{D}_u$.

Therefore, to find an optimal liability rule, it suffices to consider D under which (25) holds.

■

Proof of Proposition 2. (1) Suppose that there exists $D^* = (D_o^*, D_u^*)$ such that both (i) and (ii) in (27) hold. Then, since $\Phi_M^* = \theta\alpha_o D_o^*$ and $\Phi_m^* = (1-\theta)\alpha_u D_u^*$ from (i), the expert will choose treatment efficiently if he either learns t or chooses among $\{(N, T_M), (N, T_m)\}$ without knowing t. Moreover, from (ii) we have $\hat{k}(D^*, \Phi^*) = \overline{\Phi}_{D^*} = k^*$, which ensures that the expert will also choose his diagnostic effort efficiently. Therefore full efficiency is achieved under D^*. Notice that conditions (6) and (16) are both satisfied under D^*.

Next, we show that if the model has an equilibrium that achieves full efficiency, then there must exist some efficient D^* under which both (i) and (ii) in (27) hold. Suppose, to the contrary, that the model has an equilibrium that achieves full efficiency but at any efficient liability D^* either (i) or (ii) is violated. If (i) is violated, then $\overline{\Phi}^* < \overline{\Phi}_{D^*}$ and either $\Phi_M^* < \theta\alpha_o D_o^*$ or $\Phi_m^* < (1-\theta)\alpha_u D_u^*$. But from Lemma 5, if the model has an equilibrium that achieves full efficiency under D^*, there must also exist an efficient D^* that satisfies (25), which is a contradiction. On the other hand, if (i) holds but (ii) is violated, then $\hat{k}(D^*, \Phi^*) = \overline{\Phi}_{D^*} \neq k^*$, which means that under D^* the expert's choice of diagnostic effort is not efficient, again a contradiction.

(2) Suppose full efficiency can be achieved under some D^* with $D_o^* > 0$ and $D_u^* > 0$, then $\Phi_M^* = \theta\alpha_o D_o^* > 0$, $\Phi_m^* = (1-\theta)\alpha_u D_u^* > 0$, and $\overline{\Phi}^* = \overline{\Phi}_{D^*} = k^*$. Now consider $\tilde{D}^*$ with $\tilde{D}_o^* = \frac{k^*}{\theta(1-\theta)\alpha_o}$ and $\tilde{D}_u^* = 0$. Under $\tilde{D}^*$, from Lemma 4 it is optimal for the consumer to set $\tilde{\Phi}_M^* = \theta\alpha_o\tilde{D}_o^* = \frac{k^*}{1-\theta}$ and $\tilde{\Phi}_m^* = (1-\theta)\alpha_u\tilde{D}_u^* = 0$,[33] under which the expert would make the same choice as under D^* when he learns t or when he must choose among $\{(N, T_M), (N, T_m)\}$. Moreover, because $\overline{\Phi}_{\tilde{D}^*} = \overline{\Phi}_{D^*} = k^*$, the expert will also make the same efficient diagnostic choice under $\tilde{D}^*$. Therefore $\tilde{D}^*$ is also efficient. A similar argument shows that $\tilde{D}^*$ with $\tilde{D}_o^* = 0$ and $\tilde{D}_u^* = \frac{k^*}{\theta(1-\theta)\alpha_u}$ is also an efficient liability.

[33] It is important that $D_o^* > 0$ and $D_u^* > 0$ for the argument here. If, for instance, $D_o^* = 0$, then it might be the case that $\Phi_M^* = \theta\alpha_o D_o^*$ only because Φ_M^* cannot be lowered below 0. In this case, $\tilde{D}_o^* = \frac{k^*}{\theta(1-\theta)\alpha_o}$ and $\tilde{D}_u^* = 0$ may not induce $\tilde{\Phi}_M^* = \theta\alpha_o\tilde{D}_o^* = \frac{k^*}{1-\theta}$.

(3) First, suppose that $\Delta C \leq \Delta C^*$ so that $k^* = \theta(\Delta C + z_o)$ and $W(N, T_M) > W(N, T_m)$. We show that if z_u is sufficiently large, then $D^* = (D_o^*, D_u^*) = \left(\frac{k^*}{\theta(1-\theta)\alpha_o}, 0\right) = \left(\frac{\Delta C + z_o}{(1-\theta)\alpha_o}, 0\right)$ satisfies (27) and is thus an efficient liability. Under D^*, $\overline{\Phi}_{D^*} = \theta(1-\theta)(\alpha_o D_o^* + \alpha_u D_u^*) = k^*$. Our proof would be complete if we can show that $\overline{\Phi}^* = \overline{\Phi}_{D^*}$ if z_u is sufficiently large. Suppose, to the contrary, that $\overline{\Phi}^* < \overline{\Phi}_{D^*}$ or $S\left(\overline{\Phi}^*\right) > S\left(\overline{\Phi}_{D^*}\right)$. Then, because $\Phi_m^* = 0$ under $D_u^* = 0$, it must be true that $\Phi_M^* < \theta(1-\theta)\alpha_o D_o^*$. We thus have, with $\overline{C} = (1-\theta)\,C_M + \theta C_m$ and $k^* = \theta(\Delta C + z_o)$:

$$\begin{aligned} S\left(\overline{\Phi}_{D^*}\right) &= -\left[\beta + (1-\beta)F(\overline{\Phi}_{D^*})\right]\left[\overline{C} + \overline{\Phi}_{D^*}\right] - (1-\beta)\left[1 - F(\overline{\Phi}_{D^*})\right]\left[C_M + \theta z_o\right] \\ &= -\left[\beta + (1-\beta)F(k^*)\right]\left[\overline{C} + k^*\right] - (1-\beta)\left[1 - F(k^*)\right]\left[C_M + \theta z_o\right], \end{aligned}$$

and

$$\begin{aligned} S\left(\overline{\Phi}^*\right) &= \max_{\overline{\Phi} \in [0, k^*]} \left\{ -\left[\beta + (1-\beta)F(\overline{\Phi})\right]\left[\overline{C} + \overline{\Phi}\right] - (1-\beta)\left[1 - F(\overline{\Phi})\right]\left[C_m + (1-\theta)z_u\right] \right\} \\ &= -\left[\beta + (1-\beta)F(\overline{\Phi}^*)\right]\left[\overline{C} + \overline{\Phi}^*\right] - (1-\beta)\left[1 - F(\overline{\Phi}^*)\right]\left[C_m + (1-\theta)z_u\right]. \end{aligned}$$

Notice that $S\left(\overline{\Phi}_{D^*}\right)$ is independent of z_u, $\overline{\Phi}^* \leq \overline{\Phi}_D$, and $S\left(\overline{\Phi}^*\right)$ is decreasing in $C_m + (1-\theta)z_u$. Hence, if z_u is sufficiently large, we would have

$$C_m + (1-\theta)z_u > \frac{-S\left(\overline{\Phi}_{D^*}\right) - \left[\beta + (1-\beta)F(\overline{\Phi}^*)\right]\left[\overline{C} + \overline{\Phi}^*\right]}{(1-\beta)\left[1 - F(\overline{\Phi}^*)\right]}$$

and $S\left(\overline{\Phi}_{D^*}\right) > S\left(\overline{\Phi}^*\right)$, a contradiction.

Next, suppose $\Delta C > \Delta C^*$ so that $k^* = (1-\theta)(z_u - \Delta C)$ and $S_{NM} < S_{Nm}$. We show that $D^* = \left(0, \frac{k^*}{\theta(1-\theta)\alpha_u}\right) = \left(0, \frac{z_u - \Delta C}{\theta \alpha_u}\right)$ achieves full efficiency in equilibrium if z_o is sufficiently large. We have:

$$\begin{aligned} S\left(\overline{\Phi}_{D^*}\right) &= -\left[\beta + (1-\beta)F(\overline{\Phi}_{D^*})\right]\left[\overline{C} + \overline{\Phi}_{D^*}\right] - (1-\beta)\left[1 - F(\overline{\Phi}_{D^*})\right]\left[C_m + (1-\theta)z_u\right] \\ &= -\left[\beta + (1-\beta)F(k^*)\right]\left[\overline{C} + k^*\right] - (1-\beta)\left[1 - F(k^*)\right]\left[C_m + (1-\theta)z_u\right]. \end{aligned}$$

If $\overline{\Phi}^* = \arg\max S\left(\overline{\Phi}\right) < \overline{\Phi}_{D^*}$, which implies that $\Phi^*_m < \theta(1-\theta)\alpha_u D^*_u$,

$$S\left(\overline{\Phi}^*\right) = -\left[\beta + (1-\beta)F(\overline{\Phi}^*)\right]\left[\overline{C} + \overline{\Phi}^*\right] - (1-\beta)\left[1 - F(\overline{\Phi}^*)\right][C_M + \theta z_o].$$

A similar argument as above shows that if z_o—and hence $C_M + \theta z_o$—is sufficiently large, then $S\left(\overline{\Phi}_{D^*}\right) > S\left(\overline{\Phi}^*\right)$, contradicting with $\overline{\Phi}^* = \arg\max S\left(\overline{\Phi}\right) < \overline{\Phi}_{D^*}$.

Finally, we show that there is $\hat{\beta} < 1$ such that if $\beta > \hat{\beta}$, there exists no $D^* = (D^*_o, D^*_u)$ under which both (i) and (ii) in (27) hold in equilibrium. Consider any D such that

$$\overline{\Phi}^* = \overline{\Phi}_D \equiv \theta(1-\theta)(\alpha_o D_o + \alpha_u D_u).$$

Then, from (19),

$$\begin{aligned} S\left(\overline{\Phi}^*\right) &= S\left(\overline{\Phi}_D\right) = \left[\beta + (1-\beta)F(\overline{\Phi}_D)\right] S\left(E_T\right) + (1-\beta)\left[1 - F(\overline{\Phi}_D)\right]\max\left\{S_{NM}, S_{Nm}\right\} \\ &= \left[\beta + (1-\beta)F(\overline{\Phi}_D)\right]\left(-\overline{\Phi}_D - \overline{C}\right) + (1-\beta)\left[1 - F(\overline{\Phi}_D)\right]\max\left\{-C_M - \theta z_o, -C_m - (1-\theta)z_u\right\}. \end{aligned}$$

Hence

$$\begin{aligned} S'\left(\overline{\Phi}_D\right) &= -\left[\beta + (1-\beta)F(\overline{\Phi}_D)\right] + (1-\beta)f(\overline{\Phi}_D)\left(-\overline{\Phi}_D - \overline{C}\right) \\ &\quad - (1-\beta)\, f(\overline{\Phi}_D)\max\left\{-C_M - \theta z_o, -C_m - (1-\theta)z_u\right\}, \end{aligned}$$

and $S'\left(\overline{\Phi}_D\right) \to -\beta < 0$ as $\beta \to 1$. Therefore, there exists some $\hat{\beta} < 1$ such that if $\beta > \hat{\beta}$, $S'\left(\overline{\Phi}_D = k^*\right) < 0$, and thus $\overline{\Phi}^* = \overline{\Phi}_D < k^*$. That is, if $\beta > \hat{\beta}$, there exists no $D^* = (D^*_o, D^*_u)$ under which both (i) and (ii) in (27) hold in equilibrium. ■

Proof of Corollary 1. In Lemma 5, we showed that in searching for a welfare-maximizing liability rule $D^{**} \equiv (D^{**}_o, D^{**}_u)$, we can focus on the liabilities that induce expected markup $\overline{\Phi}_{D^{**}} = \theta(1-\theta)(\alpha_o D^{**}_o + \alpha_u D^{**}_u)$ in equilibrium, with $\Phi^*_M = \theta\alpha_o D^{**}_o$ and $\Phi^*_m = (1-\theta)\alpha_u D^{**}_u$. Condition (16) implies that $\overline{\Phi}_{D^{**}} \leq k^*$. Furthermore, $\hat{k}(D^{**},\Phi^*) = \overline{\Phi}_{D^{**}}$, and welfare is given by

(26). Because

$$\begin{aligned}W'(\hat{k}) &= (1-\beta)f(\hat{k})\left[-\theta C_m - (1-\theta)C_M - \max\{-\theta z_o - C_M, -(1-\theta)z_u - C_m\} - \hat{k}\right] \\ &= (1-\beta)f(\hat{k})\left[\min\{\theta(z_o+\Delta C), (1-\theta)(z_u - \Delta C)\} - \hat{k}\right] \\ &= (1-\beta)f(\hat{k})\left(k^* - \hat{k}\right) > 0 \text{ for } \hat{k} < k^*,\end{aligned}$$

welfare increases with $\hat{k}$ for $\hat{k} < k^*$. Hence, a second-best liability D^{**} must induce the highest possible expected equilibrium markup, $\overline{\Phi}_{D^{**}}$.

At zero liability with $D_0 = (0,0)$, we have $\overline{\Phi}^* = \overline{\Phi}_{D_0}$. Suppose that the first best cannot be achieved. Then for $\hat{D}$ that satisfies $\theta(1-\theta)(\alpha_o \hat{D}_o + \alpha_u \hat{D}_u) = k^*$, we must have $\overline{\Phi}^* < \overline{\Phi}_{\hat{D}}$. Now starting from $\hat{D}$ and continuously reducing $\overline{\Phi}_D$, there must be some highest $\overline{\Phi}_{\tilde{D}}$ at which $\overline{\Phi}^* = \overline{\Phi}_{\tilde{D}}$. Then $\tilde{D}$ maximizes welfare, and hence $D^{**} = \tilde{D}$ is a second-best liability. ■

Proof of Proposition 3. When (28) holds, making use of the equilibrium markups in (29) we have

$$\overline{\Phi}^* = \theta \Phi^*_m + (1-\theta)\,\Phi^*_M = \theta\,(1-\theta)\,(\alpha_o D^*_o + \alpha_u D^*_u) = \overline{\Phi}_{D^*},$$

and

$$\overline{\Phi}_{D^*} = \theta\,(1-\theta)\left[\alpha_o \frac{k^*}{\theta\alpha_o} + \alpha_u \frac{k^*}{(1-\theta)\alpha_u}\right] = k^*.$$

Therefore, from Proposition 2, full efficiency is achieved under D^* in (30). ■

Proof of Proposition 4. Note that under (31), the expert will choose k efficiently:

$$\begin{aligned}\hat{k}(D^*, \Phi^*) &= \min\{\theta\,(\Phi^*_m - \Phi^*_M + \alpha_o D^*_o)\,, (1-\theta)\,(\Phi^*_M - \Phi^*_m + \alpha_u D^*_u)\} \\ &= \theta(1-\theta)(\alpha_o D^*_o + \alpha_u D^*_u) = k^*.\end{aligned}$$

Next, because

$$\pi(N, T_M) = \Phi^*_M - \theta\alpha_o D^*_o = \Phi^*_m - (1-\theta)\alpha_u D^*_u = \pi(N, T_m),$$

if $k > k^*$ the expert will choose (N, T_M) when $\Delta C \leq \Delta C^*$ and (N, T_m) when $\Delta C > \Delta C^*$.

Therefore, under liability rule (31), the efficient outcome in Lemma 1 is achieved if the markups in (32) are the consumer's optimal choice.

Under (Φ_M^*, Φ_m^*) in (32),

$$\begin{aligned}\Pi(D^*, \Phi^*) &= [\beta + (1-\beta)F(k^*)]\left\{\Phi_M^* + \theta\left[(1-\theta)\alpha_u D_u^* - \theta\alpha_o D_o^*\right]\right\} \\ &\quad + (1-\beta)\left[1 - F(k^*)\right](\Phi_M^* - \theta\alpha_o D_o^*) - (1-\beta)\int_0^{k^*} t dF(t) \\ &= \Phi_M^* - \theta\alpha_o D_o^* - (1-\beta)\int_0^{k^*} t dF(t) + [\beta + (1-\beta)F(k^*)]\, k^* \\ &= (1-\beta)\left[k^* - \int_0^{k^*} F(t)dt\right] - \theta\alpha_o D_o^* + \beta k^* + (1-\beta)\int_0^{k^*} F(t)dt \\ &= 0.\end{aligned}$$

Therefore, (Φ_M^*, Φ_m^*) must be optimal for the consumer under liability (31) since total welfare is maximized while the expert receives zero ex ante expected profit, with all the surplus going to the consumer. Thus, liability (31) indeed leads to the efficient outcome. ■

Proof of Proposition 5. Suppose that liability rule $D^* = (D_o^*, D_u^*)$ satisfies

$$\theta(1-\theta)(\alpha_o D_o^* + \alpha_u D_u^*) = k^*,$$

where $k^* = \min\{\theta(\Delta C + z_o),\ (1-\theta)(z_u - \Delta C)\}$. Suppose further that the expert considers markups

$$\text{(34)} \qquad \Phi_M = \theta\alpha_o D_o^* + \delta \quad \text{and} \quad \Phi_m = (1-\theta)\alpha_u D_u^* + \delta, \quad \text{for } \delta \geq 0$$

with the expected markup being

$$\overline{\Phi} = \theta\Phi_m + (1-\theta)\Phi_M = \theta(1-\theta)(\alpha_o D_o^* + \alpha_u D_u^*) + \delta.$$

Then, from (7) and (8), we have

$$\pi(N, T_M) = \delta, \quad \pi(N, T_m) = \delta, \quad \pi(E_T) = \theta(1-\theta)(\alpha_o D_o^* + \alpha_u D_u^*) + \delta - k.$$

Under D^* and the markups in (34), both (6) and (16) are satisfied, and the expert will thus provide the right treatment if he knows t; and because $\pi(N, T_M) = \pi(N, T_m)$, he will provide the treatment between T_M and T_m that maximizes $\{S(N, T_M), S(N, T_m)\}$ if he does not know t. Moreover, the expert will choose to incur k if he initially does not learn t when

$$k \leq \hat{k}(D^*, \Phi) = \theta(1-\theta)(\alpha_o D_o^* + \alpha_u D_u^*) = k^*.$$

Therefore the markups in (34) (or, equivalently, the prices) achieve the first-best welfare W^*.

Finally, the consumer's expected surplus under these prices is

$$S\left(\overline{\Phi}\right) = \left[\beta + (1-\beta)F(k^*)\right] S\left(E_T\right) + (1-\beta)\left[1 - F(k^*)\right] \max\{S(N, T_M), S(N, T_m)\},$$

which decreases in δ because from (13), (14), and (15) we have

$$S(N, T_M) = -\delta - C_M - \theta z_o, \quad S(N, T_m) = -\delta - C_m - (1-\theta) z_u, \quad S(E_T) = -k^* - \delta - \overline{C}.$$

If $\delta = 0$, we must have $S\left(\overline{\Phi}\right) > -x$ because $S\left(E_T\right) > \max\{S\left(N, T_M\right), S\left(N, T_m\right)\} \geq -x$. On the other hand, $S\left(\overline{\Phi}\right) < -x$ if δ is sufficiently large. Hence, there exists a unique δ^* such that $S\left(\overline{\Phi}\right) = -x$ if $\delta = \delta^*$. Then the expert's markups with $\delta = \delta^*$ must maximize his profit while also maximizing welfare. That is, under D^*, the expert will optimally set $\Phi_M^* = \theta \alpha_o D_o^* + \delta^*$, $\Phi_m^* = (1-\theta)\alpha_u D_u^* + \delta^*$, and choose $\hat{k}(D^*, \Phi^*) = k^*$, under which full efficiency is achieved. ■

References

[1] Alger, I., and F. Salanie , "A Theory of Fraud and Overtreatment in Expert Markets," *Journal of Economics and Management Strategy*, 15 (2006), 853–881.

[2] Arlen, J., "Economic Analysis of Medical Malpractice Liability and Its Reform," *Research Handbook on the Economics of Torts*, (Edward Elgar, Cheltenham, UK · Northampton, MA, USA., 2013), 33–68.

[3] Arlen, J., and W. B. MacLeod, "Torts, expertise, and authority: liability of physicians and managed care organizations," *Rand Journal of Economics*, 36 (2005), 494–519.

[4] Balafoutas, L., and R. Kerschbamer, "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 (2020), 100285.

[5] Bardey, D., Gromb, D., Martimort, D., and J. Pouyet, "Controlling Sellers Who Provide Advice: Regulation and Competition," *Journal of Industrial Economics*, 68 (2020), 409–444.

[6] Bester, H., and M. Dahm, "Credence Goods, Costly Diagnosis, and Subjective Evaluation," *Economic Journal*, 128 (2018), 1367–1394.

[7] Bonroy, O., S. Lemarie, and J. Tropeano, "Credence Goods, Experts and Risk Aversion," *Economics Letters*, 120 (2013), 464–467.

[8] Brown, J. P., "Toward an Economic Theory of Liability," *The Journal of Legal Studies*, 2(1973), 323–349.

[9] Brownlee, S., *Overtreated: Why Too Much Medicine is Making Us Sicker and Poorer* (New York: Bloomsbury USA, 2008).

[10] Buck, I. D., "Caring Too Much: Misapplying the False Claims Act to Target Overtreatment," *Ohio State Law Journal*, 74 (2013), 463–513.

[11] Buck, I. D., "Breaking the Fever: A New Construct for Regulating Overtreatment," *U. C. Davis Law Review*, 48 (2015), 1261–1318.

[12] Chen, Y., and X. Hua, "Ex Ante Investment, Ex Post Remedies, and Product Liability," *International Economic Review*, 53 (2012), 845–866.

[13] Chen, Y., and T. Zhang, "Entry and Welfare in Search Markets," *Economic Journal*, 128 (2018), 52–80.

[14] Chen, Y., Z. Li, and T. Zhang, "Experience Goods and Consumer Search," *American Economic Journal: Microeconomics*, (in press).

[15] Coey, D., "Physicians' Financial Incentives and Treatment Choices in Heart Attack Management," *Quantitative Economics*, 6 (2015), 703–748.

[16] Danzon, P. M., "Liability for Medical Malpractice," *Journal of Economic Perspectives*, 5 (1991), 51–69.

[17] Darby, M., and E. Karni, "Free Competition and the Optimal Amount of Fraud," *Journal of Law and Economics*, 16 (1973), 67–88.

[18] Daughety, A. F., and J. F. Reinganum, "Product Safety: Liability, R&D, and Signaling," *American Economic Review*, 85 (1995), 1187–1206.

[19] Daughety, A. F., and J. F. Reinganum, "Product Liability, Signaling, and Disclosure," *Journal of Institutional and Theoretical Economics*, 164 (2008), 106–126.

[20] Dickstein, M. J., "Physician vs. Patient Incentives in Prescription Drug Choice," Mimeo, 2016.

[21] Demougin, D., and C. Fluet, "Preponderance of Evidence," *European Economic Review*, 50 (2006), 963–976.

[22] Demougin, D., and C. Fluet, "Rules of Proof, Courts, and Incentives," *Rand Journal of Economics*, 39 (2008), 20–40.

[23] Dulleck, U., and R. Kerschbamer, "On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods," *Journal of Economic Literature*, 44 (2006), 5–42.

[24] Dulleck, U., and R. Kerschbamer, "Experts vs. Discounters: Consumers Free-Riding and Experts Withholding Advice in Markets for Credence Goods," *International Journal of Industrial Organization*, 27 (2009), 15–23.

[25] Emons, W., "Credence Goods and Fraudulent Experts," *Rand Journal of Economics*, 28 (1997), 107–119.

[26] Emons, W., "Credence Goods Monopolists," *International Journal of Industrial Organization*, 19 (2001), 375–389.

[27] Fong, Y., "When do Experts Cheat and Whom do They Target?" *Rand Journal of Economics*, 36 (2005), 113–130.

[28] Fong, Y. and T. Liu, "Liability and Reputation in Credence Goods Markets," *Economics Letters*, 166 (2018), 35–39.

[29] Fong, Y., T. Liu, and D. J. Wright, "On the Role of Verifiability and Commitment in Credence Goods Market," *International Journal of Industrial Organization*, 37 (2014), 118–129.

[30] Gruber, J., J. Kim, and D. Mayzlin, "Physician Fees and Procedure Intensity: The Case of Cesarean Delivery," *Journal of Health Economics*, 18 (1999), 473–490.

[31] Hyndman, K., and S. Ozerturk, "Consumer Information in a Market for Expert Services," *Journal of Economic Behavior and Organization*, 80 (2011), 628–640.

[32] Hilger, N., "Why Don't People Trust Experts?," *Journal of Law and Economics*, 59 (2016), 293–311.

[33] Hua, X., "Product Recall and Liability," *Journal of Law, Economics, & Organization*, 27 (2011), 113–136.

[34] Jost, P., S. Reik, and A. Ressi, "The Information Paradox in a Monopolist's Credence Goods Market," *International Journal of Industrial Organization*, 75 (2021), 102694.

[35] Liu, F., A. Rasch, M. Schwarz, and C. Waibel, "The Role of Diagnostic Ability in Markets for Expert Services," CESifo Working Paper No. 8704, 2020.

[36] Liu, T., "Credence Goods Markets with Conscientious and Selfish Experts," *International Economic Review*, 52 (2011), 227–244.

[37] Liu, T., and C. A. Ma, "Equilibrium Information in Credence Goods," mimeo, 2021.

[38] Pesendorfer, W., and A. Wolinsky, "Second Opinions and Price Competition: Inefficiency in the Market for Expert Advice," *Review of Economic Studies*, 70 (2003), 417–437.

[39] Pitchik, C., and A. Schotter, "Honesty in a Model of Strategic Information Transmission," *The American Economic Review*, 77 (1987), 1032–1036.

[40] Shavell, S., "Strict Liability vs. Negligence," *Journal of Legal Studies*, 9 (1980), 1–25.

[41] Shavell, S., "Liability for Accidents," *Handbook of Law and Economics*, Volume 1, edited by A. Mitchell Polinsky and Steven Shavell, Elsevier B.V., 2007.

[42] Simon, M. J., "Diagnoses and Medical Malpractice: A Comparison of Negligence and Strict Liability Systems," *Bell Journal of Economics*, 13 (1982), 170–180.

[43] Spier, K. E., "Product Safety, Buybacks, and the Post-Sale Duty to Warn," *Journal of Law, Economics, & Organization*, 27 (2011), 515–539.

[44] Taylor, C.R., "The Economics of Breakdowns, Checkups, and Cures," *Journal of Political Economy*, 103 (1995), 53–74.

[45] Wolinsky, A., "Competition in a Market for Informed Experts' Services," *The Rand Journal of Economics*, 24 (1993), 380–398.