Impact of Physician Payment Scheme on Diagnostic Effort and Testing

Elodie Adida[*] Tinglong Dai[†]

[*]University of California, Riverside, School of Business, Riverside, California 92521, elodieg@ucr.edu
[†]Johns Hopkins University, Carey Business School, Baltimore, Maryland 21202, dai@jhu.edu

Diagnostic errors are common and can result in serious patient harm. Making the right diagnosis often requires significant diagnostic effort. Yet most physician payment schemes are procedure-based and do not account for diagnostic effort or accuracy due to observability issues. In this paper, we develop a parsimonious model to examine the impact of a physician payment scheme on a physician's decisions to (1) exert diagnostic effort and (2) perform a confirmatory test. High effort provides an informative (though imperfect) signal of the patient's true state; the test is confirmatory in that it is a prerequisite for diagnosing a severe condition. Our model uses a two-step diagnostic process to capture the interaction between the physician's diagnostic effort and testing decisions. We show that under a fee-for-service payment scheme, the physician may view the diagnostic effort and the confirmatory test as either complementary or substitutive, depending on the additional revenue from testing. We also reveal non-monotonic properties such that a more patient-centered physician may not exert more effort or provide a more accurate diagnosis. In addition, either a flat or differentiated payment scheme may be optimal. We also show that an alternative payment scheme, under which the revenue from the confirmatory test is contingent on its result, can induce the social optimum under certain conditions. With the advent of artificial intelligence (AI) as part of the standard of care and its increasing use as a confirmatory test, our research has implications for the design of physician payment systems in light of concerns about the potential erosion of individual attention.

Key words: Diagnostic effort, diagnostic testing, physician payment, healthcare operations management, artificial intelligence.

History: This version is dated April 17, 2023

---

1. Introduction

Most U.S. patients experience at least one diagnostic error in their lifetime (National Academies 2015). Misdiagnoses, which include erroneous, missing, or unduly delayed diagnoses, affect about 1 in 20 adults each year in the U.S. (Singh et al. 2014) and contribute to the death of approximately 160,000 patients per year (Newman-Toker et al. 2013). Physicians' diagnostic processes are an important contributor to misdiagnoses (Singh et al. 2019). Correct diagnoses are more likely when physicians exert substantial diagnostic effort, which may require them to conduct "an extensive clinical record review, listen comprehensively and gather history from patients and families, ... use diagnostic decision support and other online knowledge resources, and explore the published literature in depth" (Berenson and Singh 2018, p. 1830). Such effort, often hard to observe and thus rarely used as the basis for compensating physicians (Bester and Dahm 2017; Jelovac 2001), comes at a cost to a physician, who needs to focus more attention on the patient and spend more time researching the patient's needs (Topol 2019; Trzeciak and Mazzarelli 2019). As a result, even physicians who are dedicated to providing quality care face an effort-quality tradeoff and may not be able to afford to devote substantial diagnostic effort to every patient.

Determining whether a diagnosis is correct, either immediately or retrospectively, is challenging (Jelovac 2001). For this reason, payers rarely account for diagnostic accuracy in their reimbursements to providers. Conventional payment schemes specify reimbursements largely based on what procedures have been performed. Even in emerging payment schemes, "indicators of provider performance related to making accurate and timely diagnoses, especially where diagnostic error is common, are virtually absent from current performance measure sets" (Berenson and Singh 2018, p. 1829). Mirroring the practice, the literature has paid scant attention to the impact of the physician payment scheme on diagnostic decision-making processes and corresponding diagnostic accuracy.

The use of confirmatory testing, which is frequently required before a physician can diagnose a severe condition, further complicates the picture. For example, a positive result from a computerized tomography (CT) scan is usually required before a cardiologist can diagnose a pulmonary embolism (sudden blockage in a lung artery). Due to their cost and radiation risk, CT scans should only be performed for individuals whose ex-ante risk of pulmonary embolism is high. However, considerable evidence shows CT scans are performed on low-risk patients (Abaluck et al. 2016; Alhassan et al. 2016; Kline et al. 2020). Another example of confirmatory testing is a biopsy that is often required to confirm a melanoma diagnosis. Recent evidence has emerged that dermatologists use "a lower threshold to biopsy" than necessary, leading a substantial proportion of patients to "receive no benefit but nonetheless face the harms of scarring, wound infection, out-of-pocket costs, and the prospect of frequent surveillance" (Welch et al. 2021). Because physicians usually receive additional revenue when using such tests—and these tests simplify the diagnostic process—their incentives can be misaligned with the payer's, who often bears a large portion of the test cost. The design of current payment systems does not incorporate how to influence the physician's decision to (1) exert diagnostic effort and (2) order a confirmatory test.

In this work, we study how a physician payment scheme influences diagnostic effort and testing decisions, which jointly determine the diagnostic accuracy. We develop a parsimonious model of diagnostic decision-making to understand the impact of the physician payment system on effort and testing decisions and, hence, on diagnostic accuracy. Our model of medical decision-making highlights two forces intensifying the incentive misalignment present in the diagnostic process: on the one hand, a physician may choose not to exert costly diagnostic effort, even though the effort can help the physician reach a correct diagnosis. On the other hand, financial incentives can induce the physician to use the confirmatory test even when it is not clinically indicated, incurring a high cost to the patient and to the payer.

To put our model setup in a concrete context, consider a dermatologist who examines a patient for a potential diagnosis of cutaneous melanoma. Over the last few decades, cutaneous melanoma has become the third most frequently diagnosed cancer in the U.S., but with little improvement in survival. A key factor contributing to the dramatic increase in diagnosed cutaneous melanomas is the lower clinical threshold for biopsy (Welch et al. 2021). The dermatologist typically begins with a schematic consultation, which includes a basic physical examination and a brief review of the patient's medical history to determine the patient's risk level. The dermatologist may then conduct a more in-depth consultation, which corresponds to a high effort level in our model.[1] An in-depth consultation entails acquiring "the history of the lesion, the individual's risk factors, a more extensive assessment of the whole skin of the patient" and helps the physician assess the need to perform a biopsy (Topol 2019, p. 134). Next, the dermatologist can perform a biopsy to confirm the melanoma diagnosis; not performing a biopsy means a diagnosis of absence of melanoma. If a biopsy is performed, the patient shares the cost of the procedure with the payer and is exposed to a risk of complication from wound infection and bleeding (Wahie and Lawrence 2007; Welch et al. 2021). The dermatologist receives a technical fee for ordering and interpreting the result of a biopsy (Skaggs and Coldiron 2021) and thus has a financial incentive to order it. Carr (2021) quotes Welch et al. (2021) as stating, "Every biopsy [dermatologists] take, they get extra money, and, historically, skin biopsies have paid very well."

We make the following contributions. First, we introduce a novel model of physician decision-making that includes both diagnostic effort and testing decisions for a heterogeneous patient population. We compare the physician's optimal policy under a fee-for-service payment system with that at the social optimum. We also compare a variety of performance metrics, including diagnostic accuracy, effort level, and social welfare. These analyses enable us to shed new light on whether diagnostic effort and testing substitute for or complement each other, which has eluded the medical community thus far: whereas some argue physicians can use diagnostic tests to substitute for their

[1] The high effort level can mean more time spent interacting with patients (Trzeciak and Mazzarelli 2019). It can also entail more attention and presence during a patient visit, as Topol (2019, p. 294) argues, "[P]atients want doctors to be present, with intentional listening and undivided attention. That rarely occurs now. Rather than listening, doctors interrupt. Indeed, it only takes an average of eighteen seconds from the start of an encounter before doctors interrupt their patients... This desire to cut to the chase instead of giving the patient a chance to tell her narrative certainly matches up with the extreme time pressure that doctors and clinicians are facing."

diagnostic effort (e.g., Sirovich 2011; Bertakis and Azari 2011), others (e.g., Trzeciak and Mazzarelli 2019) argue diagnostic effort and testing can be complementary, because diagnostic tests could be omitted if the physician paid more attention to certain indicators. We find the physician may view diagnostic effort and confirmatory testing as either *complementary* or *substitutive*, depending on the additional revenue received from testing.

Second, we show the physician's decision-making exhibits non-monotonicity. Intuitively, a flat payment (which pays the physician the same with or without testing) should eliminate non-clinical influences to focus solely on the patient's well-being. Yet, a flat payment scheme may not provide sufficient incentives to perform the test for some patients, due to the patient's test cost share, leading to more misdiagnoses and lower social welfare than a differentiated payment scheme. Likewise, one may anticipate that a more patient-centered physician exerts a high diagnostic effort for more patients and achieves higher diagnostic accuracy. By contrast, we show that in the pursuit of diagnostic accuracy, a more patient-centered physician may be more likely to over-utilize the test. Because more intensive testing reduces the informative value of a high level of diagnostic effort, a more patient-centered physician can thus exert a lower level of diagnostic effort for certain patients. We also show that in some cases, out of concern for the patient cost share, a more patient-centered physician may be less likely to test certain patients, resulting in lower diagnostic accuracy.

Finally, we demonstrate that an alternative physician payment scheme in which the physician's payment is contingent on the outcome of the confirmatory test can result in the social optimum when the physician's level of patient-centeredness is not very high. By compensating physicians for confirmatory testing only when the test result is positive, this incentive scheme essentially rewards the physician for using the test when necessary (which is facilitated by higher effort), and thus alleviates the tension between diagnostic effort and confirmatory testing.

Our paper is an initial attempt to understand the effect of the physician payment system on the diagnostic decision-making process. By modeling decisions pertaining to both diagnostic effort and confirmatory testing, our paper offers novel insights into how physician payment systems can induce a delicate balance between diagnostic effort and testing. In the era of artificial intelligence (AI), there is growing concern that the use of AI tools may lead to a reduction in the amount of personalized care provided to patients (Dai and Tayur 2022). Our research has implications for the design of AI-based physician payment systems, especially as AI increasingly becomes part of the standard of care and serves as a confirmatory test (Price et al. 2019). Our alternative physician payment system is broadly consistent with the New Technology Add-On Payment (NTAP) system, which provides additional payment to a hospital for the use of new technologies when significant clinical improvement, such as early diagnosis, is demonstrated (Parikh and Helmchen 2022). NTAP is now the predominant method of reimbursing hospitals for the use of AI in clinical practice.

2. Literature

Our work contributes to three strands of literature: (1) the operations-economics interface literature on diagnosis and treatment, (2) the healthcare operations literature on financial incentives, and (3) the health economics literature on diagnostic processes.

First, our paper is connected to a stream of operations-economics interfaces literature that examines medical diagnosis and treatment (e.g., Debo et al. 2008; Durbin and Iyer 2009; Paç and Veeraraghavan 2015), which builds on the credence goods literature. As reviewed by Dulleck and Kerschbamer (2006), the credence goods literature typically assumes (1) an expert can always *accurately* and *costlessly* ascertain a client's true condition and (2) the expert has an informational advantage over the client and thus may be tempted to provide unnecessary services. Several papers (Alizamir et al. 2012; Dai et al. 2017; Dai and Singh 2020) relax these assumptions by allowing the expert to be imperfect, such that the diagnostic accuracy is influenced by the intensity of testing. Our paper departs from this literature in two ways. First, in our setting, even after exerting costly, unobservable diagnostic effort, the physician can reach a misdiagnosis. Second, the expert's diagnostic process is not fully observable in our model, so moral hazard arises.

Second, a growing body of healthcare operations literature examines how to design new payment schemes for medical services to better align incentives and improve outcomes (see, e.g., Betcheva et al. 2021; Dai and Tayur 2020; Keskinocak and Savva 2020, for recent reviews). The literature has explored the impact of payment contracts between payers and providers in a variety of contexts, including dialysis for end-stage renal disease patients (Fuloria and Zenios 2001; Lee and Zenios 2012), global health (Natarajan and Swaminathan 2018), outpatient scheduling (Jiang et al. 2012), chronic care (Zorc et al. 2017), hospital readmissions (Zhang et al. 2016; Andritsos and Tang 2018), hospital-acquired conditions and quality of care for acute inpatient services (Bastani et al. 2017), provider-to-provider referral contracts (Adida and Bravo 2019), and the role of competition (Jiang et al. 2020). The literature has also investigated the impact of alternative physician payment schemes, such as reference pricing (Nassiri et al. 2022) and bundled payments (Adida et al. 2017; Andritsos and Tang 2018; Vlachy et al. 2023; Guo et al. 2019).To our knowledge, our paper is the first in the healthcare operations literature to investigate the impact of the payment scheme on physician decision-making leading to *diagnosis*. Our findings broaden the scope of this literature.

An emerging theme in the healthcare operations management literature revolves around the use of AI in day-to-day healthcare workflows (Dai and Tayur 2022). Although our model is agnostic about whether the confirmatory test is based on conventional technology or AI, it has important implications for the design of payment schemes in which physicians are required to use AI before making a final diagnosis. In this regard, our paper joins several recent papers (see, e.g., Dai and Singh 2022; de Verícourt and Gurkan 2023; Mullainathan and Obermeyer 2021; Orfanoudaki et al.

2022) in deepening the field's understanding of the implications of AI in healthcare delivery and health policy design.

Third, the health economics literature has studied—both empirically (e.g., Afendulis and Kessler 2007; Epstein and Johnson 2012; Clemens and Gottlieb 2014) and analytically (e.g., Jelovac 2001; Marinoso and Jelovac 2003; Allard et al. 2014; Bester and Dahm 2017)—how financial incentives can influence physicians' diagnostic decisions. Here, we briefly review several analytical modeling papers that consider the cost and unobservability of diagnostic effort, the possibility of incorrect diagnosis, and the presence of moral hazard. Focusing on primary care physicians and their role as gatekeepers, both Marinoso and Jelovac (2003) and Allard et al. (2014) consider the impact of physician compensation on diagnostic effort and referral decisions. Their models do not account for the possibility of performing a confirmatory test as we do in this paper. Jelovac (2001) obtains optimal payment contracts when physician effort and patient health status are not contractible, and the physician may have repeated interactions with the patient. Her model captures double moral hazard due to both hidden action and hidden information. She discovers that when repeated patient visits are possible, the optimal contract includes supply-side cost sharing to incentivize physician effort and adequate treatment. Bester and Dahm (2017) also capture moral hazard and repeated visits, without the presence of an insurer, and when the patient subjectively evaluates the treatment outcome, which determines payment. Their analysis supports a prospective reimbursement system with equal markups (based on expected costs). The findings of the latter two papers hinge on the possibility of a repeated visit in the case of an erroneous initial diagnosis and would not hold without repeated visits. Moreover, different from these two papers, our paper considers the case in which the physician must decide whether to use a costly confirmatory test, a distinguishing feature that allows us to shed new light on the interaction between effort and testing decisions.

3. Model

We describe our model setup in Section 3.1. We then discuss modeling assumptions in Section 3.2.

3.1. Model Description

A patient (hereafter "he") visits a physician (hereafter "she") to seek a diagnosis with regard to a medical condition. The patient's true state, denoted by $s \in \{\underline{s}, \bar{s}\}$, can be either mild ($s = \underline{s}$), or severe ($s = \bar{s}$). After a schematic consultation (e.g., a basic physical exam), the physician estimates the prior likelihood that the patient suffers from the severe condition. This prior is denoted by $p \in (0,1)$ (i.e., $\Pr(s = \bar{s}) = p$). We model the prior as being drawn from a probability distribution with support $[0,1]$, mean μ, probability density function $f(\cdot)$, and cumulative distribution function $F(\cdot)$. The prior p is the source of patient heterogeneity, and for a given patient encounter, the physician relies on her estimation of p to determine her diagnostic decisions, as described next. Although, in reality, patient heterogeneity may derive from other sources (e.g., cost share of the test), we focus on clinical characteristics as the source of patient heterogeneity, because the physician is most likely to focus on her clinical observations of patients to differentiate her diagnostic decisions.

The physician's first decision in her encounter with the patient is the effort level.[2] Namely, after estimating the prior via a basic exam, the physician may choose to either exert a high effort level via a more thorough exam or not. We denote by $e \in \{L, H\}$ the effort level, where a low effort level ($e = L$) means the physician only performs a basic exam, whereas a high effort level ($e = H$) means the physician spends more time to more thoroughly assess the patient's condition. The physician incurs cost c^e when exerting high effort; without loss of generality, we normalize the cost of exerting low effort to zero. In line with the literature (e.g., Lien et al. 2004; Eggleston 2005; Allard et al. 2014; Bester and Dahm 2017; Andritsos and Tang 2018; Adida and Bravo 2019), we consider the effort to be unobservable to the payer and non-reimbursable. The cost of effort can reflect the physician's opportunity cost due to the need to spend additional time with the patient or on the patient's case (Trzeciak and Mazzarelli 2019); it can also include the mental load due to more attention and focus (Topol 2019).

If the physician exerts low effort, the patient's probability of having a severe condition is the prior p, as learned from the basic exam. Stated differently, the basic exam provides an unbiased prior of the patient's condition. This assumption is consistent with the health economics literature that models how the level of diagnostic effort affects diagnostic accuracy (see, e.g., Bester and Dahm 2017; Jelovac 2001). If the physician exerts high effort, the extra time spent with the patient generates a private signal $\sigma \in \{\underline{\sigma}, \bar{\sigma}\}$. A signal $\underline{\sigma}$ is not indicative of a severe condition, whereas a signal $\bar{\sigma}$ is indicative of a severe condition. The signal precision, denoted by θ, represents the probability that the signal matches the true patient condition; specifically,

$$\Pr(\sigma = \underline{\sigma} | s = \underline{s}) = \Pr(\sigma = \bar{\sigma} | s = \bar{s}) = \theta.$$

We assume $1/2 < \theta < 1$ such that this signal is informative but imperfect.

The physician's second decision is whether to perform a confirmatory test, a decision we denote by $t \in \{0, 1\}$. The diagnostic test is *confirmatory* in the sense that the physician cannot diagnose

[2] Modeling non-reimbursable effort as a provider's decision is common in the literature. In the health economics literature, Lien et al. (2004) and Eggleston (2005) model the provider as deciding the effort level, which is costly and has an impact on the quality of the service provided to the patient but has no direct impact on the revenue. In the healthcare operations literature, Andritsos and Tang (2018) assume the provider selects an effort level that affects the patient's chance of readmission but does not trigger a reimbursement; Adida and Bravo (2019) model both preventive and treatment efforts that have an effect on health outcomes in the interaction between two providers, without generating any reimbursement.

a severe condition without performing the test. Examples of confirmatory tests include the computerized tomography (CT) scan, which is a standard tool for diagnosing pulmonary embolism, and a skin biopsy, which is a definitive test to confirm melanoma. In addition, as AI-enabled diagnostic tests become more accurate, they are increasingly being incorporated into the standard of care and used as confirmatory tests alongside conventional tests (Price et al. 2019). If $t=1$, a test is performed and its result is either positive (consistent with a severe condition) or negative (consistent with a mild condition). If $t=0$, no test is performed and the diagnosis must be that the patient's condition is mild. For simplicity of analysis, we assume the test is perfect and thus always reveals the patient's true condition. Hence, when a test is performed, the result is positive whenever $s=\bar{s}$ and negative whenever $s=\underline{s}$. Therefore, with testing, the physician's diagnosis is always correct because it matches the test result. However, in the absence of testing, the diagnosis (which is necessarily that of a mild condition) could be erroneous.

The test is costly for both the patient and the payer. We denote by C^t the total cost of the test (including the patient's and the payer's shares). This total cost includes not only the financial cost, but also any non-financial cost associated with undergoing the test. The patient's cost share is denoted by c^t; the payer's share is thus given by $C^t - c^t$. The cost of testing to the patient (c^t) includes both a financial component (e.g., co-payment or co-insurance) and a non-financial component (e.g., side effects or risks of the test for the patient, such as discomfort, risk of infection, pain, scarring, and exposure to radiation). For example, a CT scan introduces "significant health risks and financial costs" (Abaluck et al. 2016, p. 3734).

Under a fee-for-service payment system, the physician receives a compensation of r^t for ordering a test (e.g., due to the time, effort, and expertise involved in handling the specimen and interpreting test results), and r^n if no test is ordered. To keep our analytical results as general as possible and for completeness of the analysis, we are agnostic regarding how r^t compares with r^n. In a practical fee-for-service context, however, additional care usually triggers a higher reimbursement level. For this reason, we focus some of our discussions on the more realistic case of $r^t \geq r^n$.

The patient's utility function U_{patient} comprises up to two parts (see Table 1). First, the patient may incur a cost c^t when a diagnostic test is used, corresponding to his cost share of the test. Second, the patient receives a benefit or a penalty according to how the diagnosis matches his true state. If the patient truly suffers from a severe condition and is (correctly) diagnosed accordingly, he receives a utility b due to receiving a correct diagnosis. If the patient truly has a severe condition and is (incorrectly) diagnosed with a mild condition ("type II error" due to the physician opting out of the diagnostic test), the patient receives a negative payoff of $(-h)$ that represents the harm from the misdiagnosis. If the patient truly suffers from a mild condition, he is always (correctly) diagnosed, because a test is required to diagnose a severe diagnosis and the test is perfect. In the case of a correct mild diagnosis, without loss of generality, we normalize the patient's benefit to zero (essentially, b and h represent the incremental benefit/penalty compared with a correct diagnosis of a mild condition). We assume $b - c^t > -h$ so that the patient's benefit from a true positive diagnosis outweighs his share of the cost of testing.

Table 1 **Patient utility as a function of the patient's true state and the physician's testing decision**

	The physician's testing decision (t)	
The patient's true state (s)	$t = 0$ (no testing)	$t = 1$ (testing)
$s = \underline{s}$ (mild condition)	0	$-c^t$
$s = \bar{s}$ (severe condition)	$-h$	$b - c^t$

To derive the patient's expected utility, we determine the probability of the two possible true patient states, which depends on the physician's effort, e. When the physician exerts low effort ($e = L$), the patient's probability of suffering from a severe condition is p. When the physician exerts high effort ($e = H$), she observes a private signal σ. Using Bayesian updating, this signal helps refine the probability that the patient suffers from a severe condition from p to $(1-\theta)p/[(1-\theta)p+\theta(1-p)]$ when the signal is not indicative of the severe condition ($\sigma = \underline{\sigma}$), and to $\theta p/[\theta p + (1-\theta)(1-p)]$ when the signal is indicative of the severe condition ($\sigma = \bar{\sigma}$).[3] Thus, the expected patient utility is as shown in Table 2 (the expectation is taken with respect to the possible true patient states).

Table 2 **Patient expected utility $\mathbb{E}[U_{\text{patient}}]$ depending on the physician's diagnostic effort decision, signal (if applicable), and testing decision**

Physician's effort decision (e)	Physician's private signal (σ)	Physician's testing decision (t)	
		$t = 0$ (no testing)	$t = 1$ (testing)
Low effort ($e = L$)	No signal	$-ph$	$pb - c^t$
High effort ($e = H$)	$\sigma = \underline{\sigma}$ (non-indicative signal)	$\frac{(1-\theta)p}{(1-\theta)p+\theta(1-p)} \cdot (-h)$	$\frac{(1-\theta)p}{(1-\theta)p+\theta(1-p)} b - c^t$
High effort ($e = H$)	$\sigma = \bar{\sigma}$ (indicative signal)	$\frac{\theta p}{\theta p+(1-\theta)(1-p)} \cdot (-h)$	$\frac{\theta p}{\theta p+(1-\theta)(1-p)} b - c^t$

The physician's objective function $U_{\text{physician}}$ comprises two parts. The first part is the physician's direct payoff. The physician is subject to financial incentives, which include the payment r^t or r^n, depending on whether a diagnostic test was used, and possibly the cost of effort c^e if the physician chooses to exert a high diagnostic effort. We denote by $\Pi_{\text{physician}}$ the resulting physician's payoff, capturing the direct financial impact of the physician's decision and shown in Table 3.

The physician is driven not only by financial incentives, but also by a concern for the patient. Hence, the physician's utility includes a second part linked to the expected patient utility:

$$U_{\text{physician}} = \Pi_{\text{physician}} + \delta \cdot \mathbb{E}[U_{\text{patient}}].$$

[3] Our model of the physician's diagnostic decision-making process builds on the literature on information gathering (e.g., Smith and Ulu 2017). Similar Bayesian updating models have been used to investigate incentive provision for information-gathering agents (e.g., Gromb and Martimort 2007).

Table 3 **Physician payoff $\Pi_{\text{physician}}$ as a function of effort and testing decisions**

	Physician's testing decision (t)	
Physician's effort decision (e)	$t=0$ (no testing)	$t=1$ (testing)
$e=L$ (low effort)	r^n	r^t
$e=H$ (high effort)	r^n-c^e	r^t-c^e

Parameter δ, which we refer to as the physician's degree of patient-centeredness, can be influenced by a variety of factors, including the physician's awareness of the financial cost borne by the patient, the perceived likelihood of being held liable in the event of a misdiagnosis, the degree of altruism, the effect of a possible reputational loss, and/or the degree of accountability (i.e., the chance of hearing about a potential misdiagnosis). In our model, physicians are considered homogeneous relative to the value of parameter δ. To maintain tractability and focus on the first-order effect of the payment scheme, we evaluate the effect of the compensation on an "average physician." The effect of possible physician heterogeneity in the value of δ on the performance of the physician payment scheme is beyond the scope of this paper and is thus left as a future research direction.

As a tie-breaking rule for the testing decision, in cases in which the physician is indifferent between ordering and not ordering a test, she opts not to. Similarly, in the event of a tie in the effort decision, the physician chooses to exert low effort.

The sequence of events is as follows: in the first stage, the physician uses a basic medical exam to estimate the patient's prior probability of suffering from the condition (p). Based on the prior, the physician chooses whether to exert a high effort level or not (i.e., whether $e=H$ or L). In the second stage, given e, and if applicable (i.e., if $e=H$), after observing the signal $\sigma\in\{\underline{\sigma},\bar{\sigma}\}$, the physician makes the testing decision $t\in\{0,1\}$. If a confirmatory test is ordered, the diagnosis matches the test result. If no test is ordered, the diagnosis is that of a mild condition.

We model social welfare as the sum of the patient's, physician's, and payer's expected payoffs:

$$SW=\Pi_{\text{payer}}+\Pi_{\text{physician}}+\mathbb{E}[U_{\text{patient}}],$$

where the payer's payoff is $\Pi_{\text{payer}}=-r^n$ when $t=0$ and $\Pi_{\text{payer}}=-r^t-(C^t-c^t)$ when $t=1$. Social welfare thus encompasses all costs imposed on the system (cost of high effort and total cost of testing), as well as the benefit/penalty imposed on the patient as a result of (in)correct diagnosis, but excludes payment flows within the system (e.g., reimbursement from payer to physician).

3.2. Discussion of Modeling Assumptions

Our model makes several simplifying assumptions aimed at reflecting the key drivers of physician decisions, while maintaining tractability. Some remarks about these assumptions are in order.

First, we model the physician's objective as maximizing a weighted sum of her direct financial payoff (compensation from the payer net of effort cost) and patient utility. The medical literature has established that physicians respond to financial incentives (e.g., Clemens and Gottlieb

2014). Moreover, considering financial gain in physicians' decision-making is common in both the healthcare operations management (e.g., Adida et al. 2017; Guo et al. 2019) and health economics (e.g., Jelovac 2001; Bester and Dahm 2017) literature. Yet, research suggests physicians are not solely motivated by financial gains and that the patient's well-being has an impact on their diagnostic decisions (e.g., Ellis and McGuire 1986; Newhouse 2004). Our model captures this effect by including a term proportional to the patient's utility in the physician's objective. Through this component, the physician is penalized when (1) a diagnostic error occurs or (2) when the patient's cost share is excessively high relative to the patient's risk level, which, in turn, incentivizes the physician to make correct diagnoses while avoiding unnecessary testing.

Second, we model the physician's compensation as linked to the testing decision. This assumption is in line with a fee-for-service payment system. Some tests may be done in the physician's office, and thus represent a new CPT code that the physician can get reimbursed for. Even if the test is not performed in the office, it may require the physician to collect a specimen to be sent to a lab (e.g., biopsy), and the physician can also bill for the specimen collection, handling, and shipping. Furthermore, testing can be the basis for categorizing the visit as at a higher complexity level, which can give rise to a higher reimbursement level (Hollmann et al. 2021).

Third, the payer's goal is to maximize social welfare, which includes the utility of the patient, the payer's payoff, and the physician's payoff. This approach is consistent with the literature, which generally takes the perspective of a public payer, such as Medicare, that is concerned about the patient population while also recognizing the importance of the physician's welfare in maintaining access to care (e.g., Guo et al. 2019; Nassiri et al. 2022). We use the physician's financial payoff rather than the physician's utility to determine social welfare to avoid double counting the patient utility (the physician's utility incorporates the patient utility multiplied by factor δ). In Appendix A3, we investigate an alternate definition of social welfare that includes the physician's utility rather than the physician's payoff, and our results continue to hold qualitatively.

Fourth, consistent with the related literature, we assume the confirmatory test is perfectly accurate. For example, Maillart et al. (2008) assume that when a screening mammogram yields an abnormal result, a perfect confirmatory test determines the patient's condition. Ayer et al. (2012) assume an abnormal mammogram screening result triggers a perfect followup test. In Hajjar and Alagoz (2022), a perfect confirmatory test is required before the condition can be diagnosed. Other examples include a diagnostic test for patients with dizziness/vertigo in an emergency department that can "identify more than 99% of strokes" (Newman-Toker et al. 2013), and genetic testing technologies that identify the genetic material (DNA or RNA) of disease-causing pathogens (e.g., SARS-CoV-2, the coronavirus underlying COVID-19) with nearly perfect accuracy (Bish et al. 2022). Finally, in the case of a CT scan, "a positive test is almost always followed up with immediate treatment" (Abaluck et al. 2016).

4. Physician Decisions

We analyze the physician's diagnostic effort and testing decisions in a fee-for-service environment in Section 4.1. We then characterize the physician's decisions at the social optimum in Section 4.2.

4.1. Optimal Policy under a Fee-for-Service Payment System

We characterize the physician's optimal diagnostic policy under a fee-for-service payment system. To do so, we use backwards induction to derive the physician's optimal decision regarding testing in Section 4.1.1 and regarding diagnostic effort in Section 4.1.2.

4.1.1. Second-Stage Decision: Confirmatory Testing. We analyze the physician's testing decision, occurring after the diagnostic effort decision. The next lemma characterizes the physician's optimal testing decision in the case of low effort. We denote $\Delta r \triangleq r^t - r^n$.

LEMMA 1. *Under a low diagnostic effort level (i.e.,* $e = L$*),*

(i) if $\Delta r \leq \delta(c^t - b - h)$*, the physician does not order a test for any patient;*

(ii) if $\delta(c^t - b - h) < \Delta r < \delta c^t$*, the physician orders a test if and only if* $p > \frac{\delta c^t - \Delta r}{\delta(b+h)}$*;*

(iii) if $\Delta r \geq \delta c^t$*, the physician orders a test for all patients.*

Lemma 1 states that when the financial incentives are sufficiently strong, their effect dominates the physician's decision-making. Indeed, when the revenue from testing is sufficiently low, the physician abstains from ordering it. However, when the revenue from testing is sufficiently high, the physician tests all patients, regardless of the patient's prior. When the revenue from testing is intermediate, the physician tests some but not all patients (i.e., those with a high-enough prior; all else being the same, the more financially lucrative the test is, the lower the threshold for testing).

Next, we examine the case in which the physician has exerted a high effort level.

LEMMA 2. *Under a high diagnostic effort level (i.e.,* $e = H$*),*

(i) if $\Delta r \leq \delta(c^t - b - h)$*, the physician does not order a test for any patient;*

(ii) if $\delta(c^t - b - h) < \Delta r < \delta c^t$*, the physician's testing decision depends on the private signal:*

(a) If the private signal is indicative (i.e., $\sigma = \bar{\sigma}$*), the physician orders a test if and only if* $p > \frac{(1-\theta)(\delta c^t - \Delta r)}{(1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t)+\Delta r]}$*;*

(b) If the private signal is non-indicative (i.e., $\sigma = \underline{\sigma}$*), the physician orders a test if and only if* $p > \frac{\theta(\delta c^t - \Delta r)}{\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t)+\Delta r]}$*;*

(iii) if $\Delta r \geq \delta c^t$*, the physician orders a test for all patients.*

Moreover, if $\delta(c^t - b - h) < \Delta r < \delta c^t$*, we have*

$$\frac{(1-\theta)(\delta c^t - \Delta r)}{(1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t) + \Delta r]} < \frac{\delta c^t - \Delta r}{\delta(b+h)} < \frac{\theta(\delta c^t - \Delta r)}{\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t) + \Delta r]}.$$

The intuition behind Lemma 2 is similar to that of Lemma 1, with the distinction that, with a signal available due to the high effort exerted in the first stage, the physician makes use of the signal when the test compensation is intermediate. In this region, an indicative signal lowers the prior threshold for ordering a test (compared with the threshold when no signal is available), whereas a non-indicative signal raises this threshold. Under a strong financial incentive, the signal is not being used because the physician's decision does not take into account the patient's prior.

4.1.2. First-Stage Decision: Diagnostic Effort. At this stage, we determine the physician's optimal effort decision, anticipating how the exerted effort will influence the availability of a signal and the testing decision in the next stage. For ease of exposition, we define

$$\bar{c}^F \triangleq (2\theta-1)(\delta c^t - \Delta r)\left[1 - \frac{\delta c^t - \Delta r}{\delta(b+h)}\right]$$
$$p_1^F \triangleq \frac{(1-\theta)(\delta c^t - \Delta r) + c^e}{(1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t) + \Delta r]}$$
$$p_2^F \triangleq \frac{\theta(\delta c^t - \Delta r) - c^e}{\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t) + \Delta r]}.$$

LEMMA 3. *In the first stage, given the patient's prior p, the physician's optimal effort decision is as follows:*

(i) If $\Delta r \le \delta(c^t - b - h)$ or $\Delta r \ge \delta c^t$ or $c^e > \bar{c}^F$, the physician exerts a low effort level for all patients.

(ii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ and $c^e \le \bar{c}^F$, the physician exerts a high effort level if and only if $p_1^F < p < p_2^F$.

Lemma 3 states that under extreme financial incentives, the physician exerts low diagnostic effort, because the testing decision will be the same for all patients regardless of the physician's effort. In addition, when the cost of effort is very high, it acts as a deterrent and the physician will refrain from exerting high effort regardless of patient characteristics. When the revenue from testing is moderate and the effort cost is not excessive, the physician chooses to exert high effort for patients with a borderline prior. In essence, for low-risk patients, the physician knows no testing is needed, whereas for high-risk patients, testing is clearly necessary, so exerting high effort (and incurring the associated cost) is unnecessary in either case. However, for borderline patients, the physician exerts high effort to elicit a signal that will assist in determining whether a test is necessary.

Combining Lemmas 1 to 3, we next characterize the physician's optimal effort level.

PROPOSITION 1. *The physician's optimal policy is as follows:*

(i) If $\Delta r \le \delta(c^t - b - h)$, the physician exerts low effort and does not order a test for any patient.

(ii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ and $c^e > \bar{c}^F$, the physician exerts low effort for all patients, and orders a test if and only if $p \ge \frac{\delta c^t - \Delta r}{\delta(b+h)}$.

(iii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ and $c^e \leq \bar{c}^F$[4], the optimal policy depends on the patient's prior:

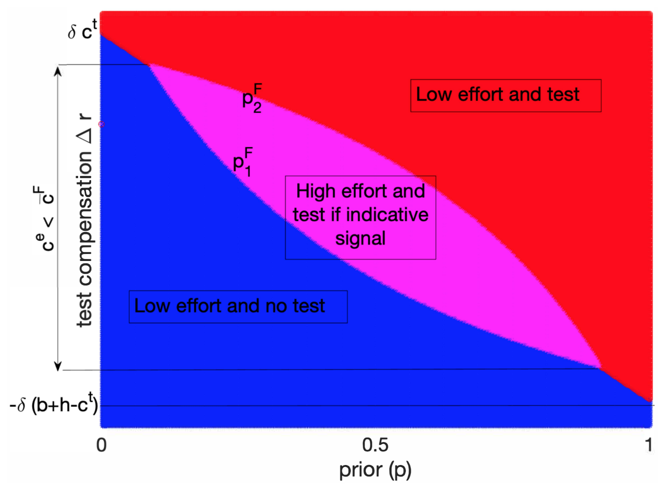
(a) if $p \leq p_1^F$, the physician exerts low effort and does not order a test;

(b) if $p_1^F < p \leq p_2^F$, the physician exerts high effort and tests according to the signal obtained (i.e., if the signal is indicative, the physician orders a test; if the signal is non-indicative, the physician does not order a test);

(c) if $p > p_2^F$, the physician exerts low effort and orders a test.

(iv) If $\Delta r \geq \delta c^t$, the physician exerts low effort and orders a test for all patients.

Figure 1 **Illustration of the physician's optimal effort and testing policy.**



Proposition 1 can be interpreted as follows. When the revenue from testing is excessively low (case (i)), the test is so financially detrimental that the physician does not test any patients. She also does not exert high effort, because the signal would not influence the testing decision. When the revenue from testing is moderate and the cost of effort is high (case (ii)), the physician exerts low effort for all patients, due to the high cost of effort, but orders a test for patients with a high prior. When the revenue from testing is intermediate and the cost of effort is low (case (iii)), three categories of patients arise. For low-risk patients, the physician makes low effort and refrains from testing, because of the low likelihood of a severe condition. For high-risk patients, conversely, the physician exerts low effort and orders a test, because of the high likelihood of a severe condition. For borderline patients, the physician exerts high effort to better assess whether a test is warranted, and the testing decision is then consistent with the signal. Finally, when the revenue from testing is high (case (iv)), this financial incentive induces the physician to test all patients. High effort is thus unnecessary, because the signal would not influence the testing decision.

[4] We show in Corollary 1 that the condition $c^e \leq \bar{c}^F$ is equivalent to Δr between two bounds, as indicated in Figure 1.

4.2. Benchmark: Social Optimum

In this section, we characterize the social optimum as a benchmark. At the social optimum, the goal is to maximize social welfare, which, as defined in Section 3.1, is the sum of the expected patient utility $\mathbb{E}[U_{\text{patient}}]$, payer's payoff Π_{payer}, and physician's payoff $\Pi_{\text{physician}}$. Because payments within the system cancel each other out, social welfare may include the total test cost, the cost of effort, and the benefit/penalty experienced by the patient from diagnosis (mis)accuracy. As a result, social welfare coincides with the physician's utility after replacing δ with 1 and replacing Δr with the payer's cost share of the test, $-(C^t - c^t)$. We thus obtain the socially optimal policy by adapting the results of Proposition 1. Let

$$\bar{c}^S \triangleq (2\theta-1)C^t\left(1-\frac{C^t}{b+h}\right)$$
$$p_1^S \triangleq \frac{(1-\theta)C^t + c^e}{(1-\theta)C^t + \theta(b+h-C^t)}$$
$$p_2^S \triangleq \frac{\theta C^t - c^e}{\theta C^t + (1-\theta)(b+h-C^t)}.$$

Proposition 2. *The socially optimal policy is as follows:*

(i) If $b+h \le C^t$, the physician exerts low effort and does not order a test for any patient.

(ii) If $b+h > C^t$ and $c^e > \bar{c}^S$, the physician exerts low effort for all patients and orders a test if and only if $p \ge C^t/(b+h)$.

(iii) If $b+h > C^t$ and $c^e \le \bar{c}^S$, the socially optimal policy depends on the patient's prior:

(a) if $p \le p_1^S$, the physician exerts low effort and does not order a test;

(b) if $p_1^S < p \le p_2^S$, the physician exerts high effort and follows the signal (i.e., if the signal is positive, the physician orders a test; if the signal is negative, the physician does not);

(c) if $p > p_2^S$, the physician exerts low effort and orders a test.

The socially optimal diagnostic strategy is determined by how the total cost of testing (C^t) compares with the combined benefit and harm of diagnosis (mis)accuracy ($b+h$). Cases (i), (ii), and (iii) of Proposition 2 have an interpretation similar to Proposition 1. However, a case analogous to case (iv) of Proposition 1 (when $\Delta r \ge \delta c^t$) does not arise in Proposition 2, because $-(C^t - c^t) < c^t$. Exerting low effort and testing all patients is never socially optimal, because the system incurs a non-zero cost of the test (C^t). By contrast, under the fee-for-service payment system, the physician's optimal strategy is to exert low effort and test all patients when the revenue from testing is sufficiently high.

4.3. Case Study

We now provide a case study to illustrate our model and analysis. Consider the example of a dermatologist who diagnoses a patient for cutaneous melanoma, as described in detail in Section 1. The dermatologist begins with a schematic consultation, and may provide a more in-depth consultation, corresponding to a high level of effort. Then, she can either perform a biopsy as a confirmatory test for a melanoma, or diagnose the absence of melanoma without performing a biopsy.

Cutaneous melanoma is a rare condition. According to the U.S. National Cancer Institute (2022), the rate of new cases of cutaneous melanoma was 215 per million per year (0.0215%), and the death rate of cutaneous melanoma was 22 per million per year (0.0022%).[5] Significant risk factors for cutaneous melanoma include older age, naturally lighter skin, blue or green eyes, blonde or red hair, and a personal or familial history of skin cancer. For example, whereas the 30–39 age group has a melanoma incidence rate of approximately 150 per million for a woman and 90 per million for a man, the 80–89 age group has a melanoma incidence rate of 550 per million for a woman and 1,800 per million for a man.

We calibrate our model parameters as follows. Aires et al. (2016) and Goldsmith (2013) use a \$200 estimate for the biopsy charges. Using a \$200 estimate, and a typical copayment rate of 20%, we estimate c^t at \$40.[6] The \$200 biopsy charge includes the total financial cost of the biopsy C^t (e.g., charged by an external laboratory) as well as the physician profit margin $\Delta r = r^t - r^n$. Following Aires et al. (2016), we estimate $\Delta r = r^t - r^n = \$36$ in the base case, and we estimate the financial cost of the test at $C^t = \$164$. (In the rest of the case study, we fix the value of C^t and vary the value of Δr to focus on the impact of the physician payment scheme.) Aires et al. (2016) estimate that a delay in diagnosing a melanoma incurs a cost of $\$33,989$, which corresponds to the difference between the benefit $+b$ experienced in the case of a correct diagnosis of the presence of the severe condition and the harm $-h$ experienced in the case of a missed diagnosis of the severe condition, that is, $b + h = \$33,989$. We estimate θ at 70% (Swetter and Geller 2022).

We consider a population with a higher incidence rate than that of the overall population (0.0215%), because low-risk patients without the common risk factors are rarely referred for melanoma diagnosis. We assume the prior probability of melanoma follows a beta distribution $\text{Beta}(\alpha, \beta)$, with a probability density function of $p^{\alpha-1}(1-p)^{\beta-1}/\text{Beta}(\alpha, \beta)$ and a support of $[0, 1]$, where $\text{Beta}(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$, and $\Gamma(\cdot)$ is the gamma function such that $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$. We calibrate the parameters of the beta distribution at $\alpha = 1$ and $\beta = 600$. Under this distribution, the mean prior probability of melanoma is $\mu = \alpha/(\alpha + \beta) = 0.00166 = 0.166\%$.

Using the above parameters, Proposition 2(ii) indicates performing a biopsy for patients with an incidence rate (i.e., prior) no less than $C^t/(b + h) = 0.0048 = 0.48\%$ (i.e., 4,800 per million) is

[5] https://seer.cancer.gov/statfacts/html/melan.html, Accessed October 1, 2022

[6] We do not consider the non-financial cost of the biopsy in this case study because, according to Aires et al. (2016), complication rates from biopsies are low and such complications are minor and can be treated with inexpensive generic antibiotics.

socially optimal. This cutoff corresponds to 5.49% of all the patients. By contrast, Proposition 1 indicates the physician's optimal strategy is quite different from what is socially optimal. By analyzing a group of medical students' medical treatment choices, Godager and Wiesen (2013) find the mode of the distribution of physician altruism is close to one. In the case of $\delta = 0.95$, the physician's decision follows one of the cases (ii)–(iv) of the proposition. Specifically, if $\Delta r \geq \$38$, Proposition 1(iv) applies; that is, the physician exerts low effort and orders a test for all patients. Otherwise (i.e., $\Delta r < \$38$), we have two cases:

- Case (ii) applies if $c^e > \bar{c}^F = \$9.59$. In this case, the physician exerts low effort and tests patients with an incident rate of no less than $\frac{\delta c^t - \Delta r}{\delta(b+h)} = 0.000248 = 0.0248\%$ (i.e., 2,480 per million) in the case of $\Delta r = \$30$, which corresponds to 86.18% of all patients.
- Case (iii) applies if $c^e \leq \bar{c}^F = \$9.59$. In this case, the physician (1) exerts low effort and does not order a test if $p \leq p_1^F$, (2) exerts low effort and orders a test if $p > p_2^F$, and (3) exerts high effort and tests according to the obtained signal if $p_1^F < p \leq p_2^F$. At $c^e = \$6$ and $\Delta r = \$15$, we have $p_1^F = 0.000571 = 0.0571\%$ and $p_2^F = 0.00104 = 0.104\%$; the physician exerts high effort for 17.48% of all patients and exerts low effort and orders a test for 53.51% of all patients.

Across cases (ii)–(iv), the physician orders a test far more frequently than at the social optimum.

In practice, the biopsy rate varies across providers. According to one observed study, a significant proportion of physicians perform biopsies on nearly all patient visits (Hamid et al. 2019). This case study illustrates that understanding diagnostic behavior helps us reveal the gap between physician behavior under the fee-for-service payment scheme and the social optimum. Stated differently, a better knowledge of diagnostic behavior in response to a payment system can lead to better design of the payment scheme, and our model can be used as a building block for policymakers to evaluate and compare different physician payment schemes.

5. Performance Measures

We define three performance measures: (1) diagnostic accuracy, (2) level of diagnostic effort, and (3) social welfare. We aim to compare fee-for-service to the social optimum with regards to these measures and to determine how the revenue from testing drives the physician's diagnostic strategy.

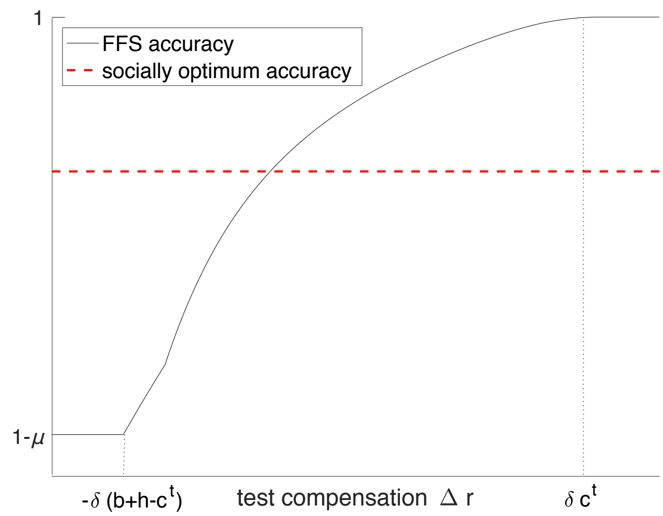
5.1. Diagnostic Accuracy

We define diagnostic accuracy as the probability that the diagnosis matches the patient condition.

Proposition 3. *The aggregate population diagnostic accuracy under fee-for-service is monotonically increasing from* $1 - \mu$ *to 1 as* Δr *increases. Moreover, the average socially optimal diagnostic accuracy is a constant equal to a value within* $[1 - \mu, 1)$.

The intuition underlying Proposition 3 is higher-powered financial incentives for testing lead to better diagnostic accuracy, due to the more frequent testing it induces (and despite the effect on the physician's diagnostic effort, which we analyze in the next section). It follows from Proposition 3 that, depending on Δr, the aggregate accuracy under fee-for-service may be less or more than at the social optimum. In other words, a threshold exists for the revenue from testing, above which the aggregate accuracy under fee-for-service is better than in the social optimum (see Figure 2).

Figure 2 **Aggregate accuracy under fee-for-service (FFS) and at the social optimum.**



5.2. Diagnostic Effort

Diagnostic effort incurs a cost to the physician, yet its benefit accrues to the patient. As such, one of the challenges in this setting is incentivizing the physician to exert an appropriate level of diagnostic effort. In this section, we compare the range of priors that lead to high diagnostic effort (when high effort is exerted on at least some patients; otherwise, the range is equal to zero) under fee-for-service and at the social optimum. We refer to "range" in the statistical sense, as the difference between the highest and the lowest priors that lead to high effort.

COROLLARY 1. *Under the fee-for-service payment scheme, the physician exerts high effort on certain patients if and only if*

$$c^e \leq (2\theta - 1)\frac{\delta(b+h)}{4} \text{ and}$$

$$\delta c^t - \delta\frac{b+h}{2}\left(1+\sqrt{1-\frac{4c^e}{\delta(b+h)(2\theta-1)}}\right) \leq \Delta r \leq \delta c^t - \delta\frac{b+h}{2}\left(1-\sqrt{1-\frac{4c^e}{\delta(b+h)(2\theta-1)}}\right).$$

For the physician to exert high effort on at least some patients, as noted in the previous section, the cost of effort cannot be too high and the revenue from testing should not be extreme (not so low that never testing is optimal, and not so high that testing everyone is optimal).

Corollary 2. *The range of priors for which the physician exerts high effort is unimodal in Δr, reaching a maximum equal to equal to $2\theta - 1 - 4c^e/(\delta(b+h))$ at $\Delta r = \delta(c^t - (b+h)/2)$.*

This result establishes that the testing revenue has a non-monotonic effect on the incentive to exert effort. As Δr first increases, more testing compensation incentivizes more effort: the extra revenue compares favorably with the patient's share of the cost (which impacts the patient-centered physician's objective). However, as Δr increases past a certain point, more compensation reduces the incentives to exert high effort: the extra revenue becomes so advantageous that the physician prefers to exert low effort and directly test more patients (see Figure 3).

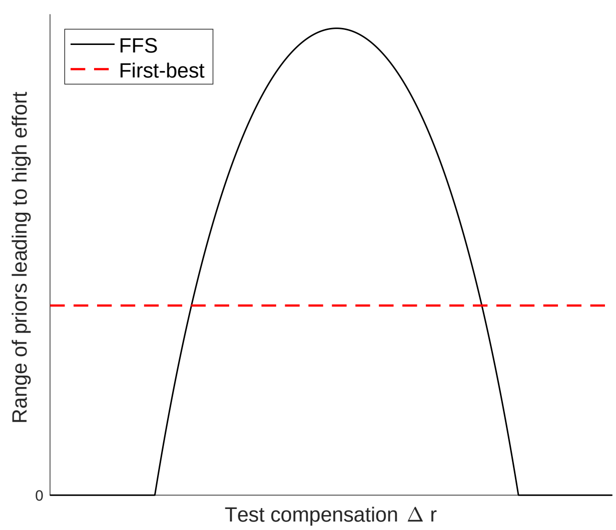
It remains to investigate how the amount of high effort exerted under fee-for-service compares with that at the social optimum.

Proposition 4. *Under fee-for-service, when Δr is such that the range of priors with high effort is at its widest, that range is wider than the socially optimal range if and only if either (i) $\delta > 1$, or (ii) $\delta \leq 1$ and either $C^t < c_0$ or $C^t > c_1$, where*

$$c_0 \triangleq \frac{b+h}{2}\left(1-\sqrt{q}\right); \qquad c_1 \triangleq \frac{b+h}{2}\left(1+\sqrt{q}\right), \text{ and } q \triangleq \frac{c^e(1-\delta)}{(2\theta-1)[c^e(2\theta-1)+\theta(1-\theta)\delta(b+h)]}.$$

Otherwise (i.e., $\delta \leq 1$ and $c_0 \leq C^t \leq c_1$), the socially optimal range of high effort is wider than the fee-for-service range of high effort for all Δr.

Figure 3 Range of priors leading to high effort under FFS and at the social optimum (with $\delta > 1$).



This result indicates the range of priors for which the physician exerts high effort (which correlates with the *number of patients* receiving high effort) is not necessarily smaller under fee-for-service than at the social optimum. Fee-for-service may lead to a wider range of priors with high effort in several scenarios: (1) when $\delta > 1$, as the physician prioritizes patient welfare over her own financial gain, whereas the social planner values both equally; (2) when $\delta \le 1$ and the total test cost is low, as the social planner prefers to directly test patients (i.e., with low effort); or (3) when $\delta \le 1$ and the total test cost is high, as the social planner prefers to avoid excessive testing, whereas high effort results in testing when the signal is positive. Figure 3 illustrates scenario (1). In all these scenarios, the physician, while acting in her own self-interest, may exert high effort for a wider range of patients under the fee-for-service payment scheme than under the social optimum. Therefore, the fee-for-service scheme does not necessarily lead to less physician effort than the socially optimal.

5.3. Social Welfare

By definition, the fee-for-service payment scheme leads to lower social welfare than in the social optimum. In this section, we analyze how varying the revenue from testing affects social welfare under fee-for-service. Namely, we address the following question: To maximize social welfare, how should testing be compensated? To answer this question, we obtain social welfare under fee-for-service in closed form (Proposition A3 in the online appendix) and we analyze how it varies with respect to Δr. The next result examines the special case of a high cost of effort.

Proposition 5. *If $c^e > (2\theta - 1)\delta(b+h)/4$, the value of Δr that maximizes social welfare under fee-for-service is*

$$\Delta r = \begin{cases} -\delta(C^t - c^t) & \text{if } b+h > C^t \\ \text{any value within } (-\infty, -\delta(b+h-c^t)] & \text{otherwise.} \end{cases}$$

Moreover, at these levels of compensation, social welfare under fee-for-service reaches the socially optimal social welfare if $b+h \le C^t$ or if $c^e > \bar{c}^S$.

This result shows that, when the cost of effort is high (such that the physician never chooses to exert high effort), to maximize social welfare, the testing compensation should be such that ordering testing would incur a penalty for the physician (rather than an added payment). If $b+h > C^t$ (i.e., testing some patients is socially optimal due to the reasonable cost of the test), this penalty is equal to the payer's cost share of the test multiplied by the parameter δ. The intent of such a compensation mechanism would be to force the physician to internalize not only the patient's cost share of the test, but also the payer's. If $b+h \le C^t$ (i.e., testing no patient is socially optimal due to the excessive cost of the test), a penalty of at least $\delta(c^t - b - h) < 0$ ensures fee-for-service also gives rise to no testing on any patient. Clearly, such a payment scheme is implausible in practice, because it amounts to charging the physician for the payer's share of the cost of a test ordered for a patient. The proof shows that if we optimize social welfare with the constraint that the testing compensation be no less than the compensation in the absence of testing, the optimal compensation is a flat payment: social welfare is maximized when the physician is paid the same regardless of whether a test is ordered (but social welfare would no longer match the social optimum).

The case $c^e \leq (2\theta - 1)\delta(b+h)/4$, where the physician exerts high diagnostic effort on certain patients, remains to be considered. Unfortunately, analytically studying the effect of Δr on social welfare in this situation is intractable. Numerically, we find social welfare appears unimodal with respect to Δr. However, the compensation scheme that maximizes social welfare is not necessarily in the negative domain, a point that we expand on in Section 6.3.

6. Managerial Implications

We now discuss managerial implications arising from our results. Section 6.1 discusses whether the physician's diagnostic effort and testing complement or substitute for each other. Section 6.2 examines the effect of patient-centeredness on several performance metrics. Section 6.3 discusses whether a flat or a differentiated payment scheme is optimal under fee-for-service. Section 6.4 investigates an alternative payment scheme and shows this scheme can yield the social optimum.

6.1. Effort and Testing: Complements or Substitutes?

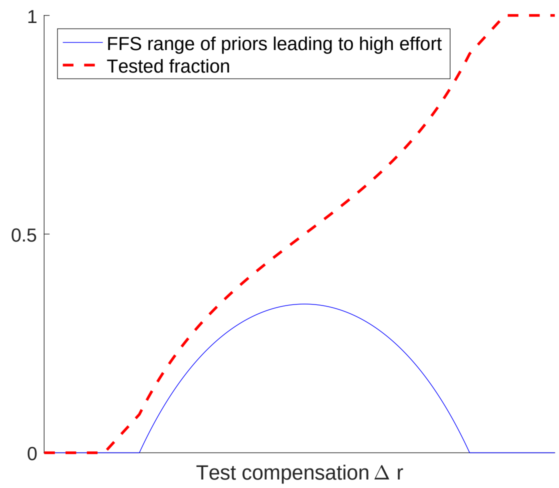
A commonly held view contends that physicians use diagnostic tests to substitute for their diagnostic effort (Sirovich 2011). Bertakis and Azari (2011) show that, in a Canadian primary care setting, when physicians paid more individual attention to patients, patients underwent significantly less testing.[7] On the other hand, Trzeciak and Mazzarelli (2019) demonstrate diagnostic tests may be omitted due to the physician's failure to pay sufficient attention to clinical indicators.

Corollary 2 (in Section 5.2) provides a more nuanced understanding of the relationship between diagnostic effort and testing, as illustrated in Figure 4. This result suggests the physician may view diagnostic effort and confirmatory testing as either complementary or substitutive, depending on the incremental revenue from testing. Specifically, when testing compensation is low (i.e., $\Delta r < \delta(c^t - (b+h)/2)$), as it increases, the physician has an incentive to exert a high diagnostic effort for certain low- or medium-risk patients, because their updated prior may trigger a test. When testing compensation is sufficiently high (i.e., $\Delta r \geq \delta(c^t - (b+h)/2)$), however, testing is so financially beneficial that the physician substitutes diagnostic effort with testing.

The above observation implies increasing the reimbursement for testing does not always mean the physician will exert less effort. Rather, in some scenarios, the physician may exert a high diagnostic effort on *more* patients in response to increased revenue from testing, which could lead to a better understanding of the patient's condition.

[7] In the case of medication prescriptions, studies have established a link between shorter visits and higher prescription rates; see Dugdale et al. (1999) for a review of these studies.

Figure 4 **Tested fraction of the population and range of priors leading to high effort under FFS.**



6.2. Effect of Patient-Centeredness

Increasing patient-centeredness has become a major focus of health leaders (Bergeson and Dean 2006). In this section, we seek to understand the effect of the degree of patient-centeredness (δ), that is, the level of awareness that the physician has of the test costs borne by patients and of the eventual health outcome. Accordingly, we investigate the effect of varying δ on three performance metrics, namely, diagnostic accuracy, the range of priors with high diagnostic effort, and social welfare. Throughout this section, we consider a fixed testing compensation $r^t - r^n \geq 0$ to focus our discussions on the most practically relevant scenarios (see Section 3.1 for a detailed discussion).

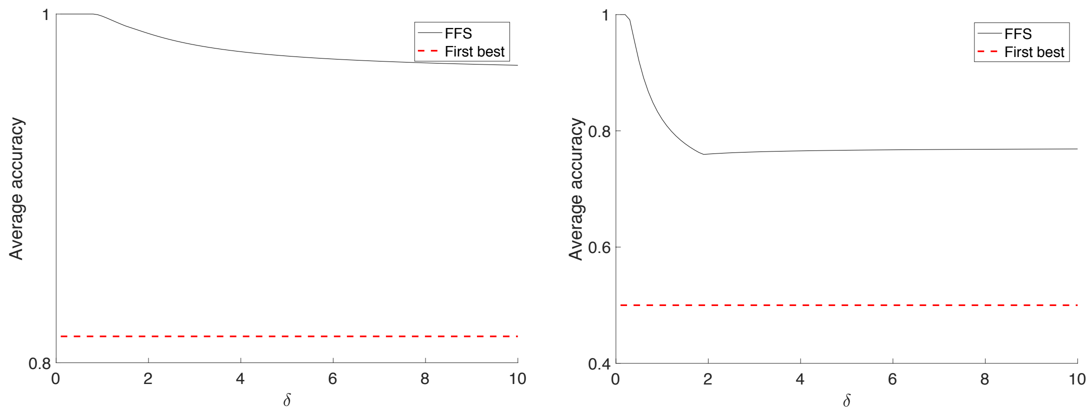
6.2.1. Effect of Patient-Centeredness on Diagnostic Accuracy. We examine how the average diagnostic accuracy varies as δ increases.

LEMMA 4. *If* $\Delta r > 0$ *and* $c^e < \theta(1-\theta)\Delta r(b+h)/((1-\theta)c^t + \theta(b+h-c^t))$, *the diagnostic accuracy is non-increasing in* δ.

Lemma 4 reveals the effect of the physician's degree of patient-centeredness on diagnostic accuracy. Interestingly, we find that for a low cost of effort and a positive revenue from testing, the accuracy *worsens* when the physician is more patient-centered. Indeed, when the patient-centeredness is low, the physician is primarily influenced by her own financial incentives, which, when $\Delta r > 0$, push for more testing and thus yield high accuracy. As patient-centeredness increases, the physician uses testing less broadly, thereby reducing accuracy, due to the rising influence of the patient's cost share of the test.

The case in which $\Delta r > 0$ and the cost of effort is high is analytically intractable. Numerically, we find that, depending on the input parameters, the accuracy is either non-increasing (as illustrated in Figure 5, left-hand-side panel) or non-monotonic (i.e., constant then decreasing then increasing, as illustrated in Figure 5, right-hand-side panel).

Figure 5 Average diagnostic accuracy under fee-for-service and at the social optimum for a varying δ.



These results show that when the provider is financially advantaged by testing, initiatives aimed at increasing the degree of patient-centeredness can worsen the diagnostic accuracy. The next result proves this issue can be remedied under certain conditions by removing financial incentives to test.

LEMMA 5. *Suppose $\Delta r = 0$. Let $D_1 \triangleq (1-\theta)c^t + \theta(b+h-c^t)$ and $D_2 \triangleq \theta c^t + (1-\theta)(b+h-c^t)$. If $f(p_1^F)/D_1^2 \geq f(p_2^F)/D_2^2$, the diagnostic accuracy is non-decreasing in δ. Otherwise, the diagnostic accuracy is unimodal (constant then increasing then decreasing) in δ.*

When payment is flat, the diagnostic accuracy may improve when the degree of patient-centeredness δ increases when the condition $f(p_1^F)/D_1^2 \geq f(p_2^F)/D_2^2$ holds or when the range of values reached by δ is not too high. Intuitively, when the payment is flat, the physician has no direct incentive to test. She is influenced by her cost of effort and the patient's utility (test cost share and benefit or harm from (in)correctness of diagnosis). As δ increases, the cost of the test pushes the physician to exert more high effort, as formally shown in the next proposition (which improves accuracy), and to refrain from unnecessary testing (which hurts accuracy); the benefit/harm from (in)correctness of diagnosis pushes the physician to test more to avoid an incorrect diagnosis. These effects interact in a non-trivial way, resulting in accuracy that may not be monotonic in δ.

6.2.2. Effect of Patient-Centeredness on Diagnostic Effort. We next focus on the effect of patient-centeredness on the range of priors for which the physician exerts high effort.

PROPOSITION 6. *If $\Delta r = 0$, the range of priors leading to high effort under fee-for-service is monotonically increasing in δ. If $\Delta r > 0$, the range of priors leading to high effort under fee-for-service is either monotonically increasing or unimodal in δ.*[8]

[8] Proposition A4 in Appendix A2 fully characterizes the conditions under which the range of priors leading to high effort is monotonic, and, when it is unimodal, for which value of δ it reaches a maximum.

Proposition 6 shows that increasing the value of δ lowers the physician's incentives to exert high effort. Although in some cases (e.g., $\Delta r = 0$), increasing δ expands the range of priors with high effort, in other cases, increasing δ could reduce this range. The reason is that when testing leads to a higher payment, increasing the degree of patient-centeredness can lead the physician to directly test more patients (accompanied by low diagnostic effort) to take advantage of the added compensation, lack of cost of effort, and fewer misdiagnoses, despite the cost of the test. This effect can be avoided using a flat payment system: when $\Delta r = 0$, increasing δ leads to an increasing range of priors with high effort, because direct testing is less advantageous to the physician, and thus, the effect of the cost of the test prompts the physician to exert high effort instead.

The following corollary provides a simple and sufficient condition for a monotonically increasing range of priors that correspond to high effort even in the case of $\Delta r > 0$.

Corollary 3. *If $0 < c^t \le (b+h)/2$, the range of priors with high effort is increasing in δ.*

Intuitively, when the cost of the test (to the patient) is sufficiently low relative to its value (i.e., $c^t \le (b+h)/2$), the test is high-value. If δ is very small, the physician directly tests every patient (with low effort). As δ increases, the cost of testing gains more importance in the physician's decision-making. Thus, to avoid unnecessary tests, she exerts high effort before testing.

6.2.3. Effect of Patient-Centeredness on Social Welfare. The next result considers the effect of patient-centeredness on the average social welfare in the case of a flat testing compensation.

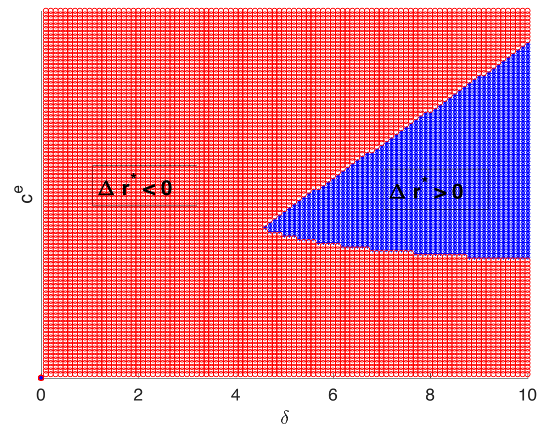
Lemma 6. *Suppose $\Delta r = 0$ and the patient priors are uniformly distributed. For δ below a threshold, social welfare is a constant with respect to δ. Above the threshold, if $c^t \ge (b+h)/2$ or if c^e is large enough, social welfare in this region is either monotonically decreasing or unimodal. Otherwise (i.e., $c^t < (b+h)/2$ and c^e is low enough[9]), social welfare in this region is either monotonically increasing or sequentially decreasing then increasing.*

This finding shows increasing patient-centeredness does not necessarily benefit social welfare, even when the revenue from testing is flat. Although this result is obtained in the special case of a uniform distribution of patient priors, its implication is more general because it establishes that the effect of patient-centeredness on social welfare is not generally monotonic.

6.3. Flat versus Differentiated Payment Schemes

We showed in Section 5.3 that, for a high cost of effort, the testing compensation that maximizes social welfare is negative. Hence, when the testing compensation is constrained to be non-negative, social welfare is highest under a flat compensation scheme (for a high cost of effort). However,

Figure 6 **Sign of the social-welfare-maximizing testing compensation for varying δ and c^e.**



this scenario may not hold true for a lower cost of effort. Indeed, we obtain numerically that the optimal testing compensation may be positive in some instances, as shown in Figure 6.

When the social planner designs the optimal payment scheme with the goal of maximizing social welfare constrained by a non-negative compensation, our study shows that if one ignores the role of physician effort in the diagnostic process (e.g., when the cost of effort is so high that the physician never chooses to exert high effort), the optimal payment scheme is flat (see Proposition 5). The finding is due to the fact that the physician does not internalize the payer's share of the cost of the test, so tends to under-count the cost of testing; as a result, higher-powered incentives for testing would only further distort the incentive structure. However, when the social planner takes into account the role of physician effort, this result may no longer hold true (see, e.g., Figure 6). A positive financial incentive for testing may be optimal to incentivize the right level of effort.

6.4. Incentive Alignment

We now analyze how to redesign the physician payment scheme to shift the physician's incentives in the direction of the social optimum. To align decisions, the goal is essentially to align the thresholds (p_1^F, p_2^F) with (p_1^S, p_2^S), as well as align $\bar{c}^F$ with $\bar{c}^S$.

With this goal in mind, we considered augmenting the current fee-for-service payment model with a combination of performance incentives and/or subsidies. For example, we considered rewarding the physician in the case of a correct severe diagnosis and/or penalizing her in the case of an incorrect mild diagnosis (assuming the outcomes could be observed and the correctness of the diagnosis identified). In the context of our model, such a performance incentive translates into inflating the value of $b+h$. Likewise, providing subsidies to the patient and/or physician for testing would boil down to adjusting parameters c^t and/or Δr. Awareness programs could help adjust the

[9] The thresholds on δ and c^e are provided in closed form in the proof in Appendix A2.

value of δ. Therefore, we examine in the next proposition whether a combination of such incentives could achieve the social optimum.

PROPOSITION 7. *Suppose b, h, c^t, δ, and/or Δr can be adjusted. The only adjustment yielding the socially optimal policy is such that $\delta = 1$ and $\Delta r = c^t - C^t$.*

This result implies that no performance incentive aiming to change the physician's reward and penalty associated with (in)correctness of diagnosis (b and h) can help obtain the social optimum. Similarly, testing subsidies for the patient and/or physician are not helpful. To align with the social optimum, δ needs to be adjusted, so the physician weighs equally her own benefit and the patient's, and the compensation scheme needs to consist of penalizing the physician for ordering a test, with a penalty equal to the payer's share of the cost of the test, so that the physician internalizes that cost and bases her decision on the total cost of the test, like the social optimum does. Because such a payment scheme lacks practicality, we next consider a different type of performance incentive.

Physicians are shown to be more likely to use diagnostic tests unnecessarily when they profit from them (Shute 2011). Motivated by this observation, we now consider a diagnosis-based payment scheme whereby, when a test is ordered, the physician's payment depends on the test result. Namely, we consider a payment scheme whereby, when the physician orders a diagnostic test, she receives a payment of r^+ if the result is positive and a payment of r^- if the result is negative. We continue to use r^n to denote the payment that the physician receives for not ordering the test. Intuitively, by differentiating the payment according to the test result, it may be possible to incentivize the physician to order a test only for those patients who are most likely suffering from the severe condition and thus have a high chance of receiving a positive test result.

The analysis of the physician's effort and testing decisions is similar to that under the fee-for-service payment scheme. For brevity of presentation, we omit the detailed analysis and summarize the physician's decision rule in the following proposition, in which we define

$$p_1^d \triangleq \frac{(1-\theta)(\delta c^t + r^n - r^-) + c^e}{(1-\theta)(\delta c^t + r^n - r^-) + \theta[\delta(b+h-c^t) + r^+ - r^n]}$$
$$p_2^d \triangleq \frac{\theta(\delta c^t + r^n - r^-) - c^e}{\theta(\delta c^t + r^n - r^-) + (1-\theta)[\delta(b+h-c^t) + r^+ - r^n]}$$
$$\bar{c}^d \triangleq \frac{(2\theta-1)(\delta c_t + r^n - r^-)[\delta(b+h-c_t) + r^+ - r^n]}{\delta(b+h) + r^+ - r^-}.$$

PROPOSITION 8. *Under the diagnosis-based payment scheme, the physician's optimal policy is as follows:*

(i) If $r^+ - r^n \leq \delta(c^t - b - h)$, the physician exerts low effort and does not order a test for any patient.

(ii) If $r^+ - r^n > \delta(c^t - b - h)$ and $c^e > \bar{c}^d$, the physician exerts low effort for all patients, and orders a test if and only if $p > \frac{\delta c^t + r^n - r^-}{\delta(b+h) + r^+ - r^-}$.

(iii) If $r^+ - r^n > \delta(c^t - b - h)$ and $c^e \leq \bar{c}^d$, the optimal policy depends on the patient's prior:

(a) if $p \leq p_1^d$, the physician exerts low effort and does not order a test;

(b) if $p_1^d < p \leq p_2^d$, the physician exerts high effort and tests according to the signal obtained (i.e., if the signal is indicative, the physician orders a test; if the signal is non-indicative, the physician does not order a test);

(c) if $p > p_2^d$, the physician exerts low effort and orders a test.

The next proposition shows this type of diagnosis-based payment scheme can align the physician's incentives with the social planner's objective.

Proposition 9. *The physician's effort and testing decisions maximize social welfare under a diagnosis-based payment scheme that satisfies*

$$b_1 \triangleq r^n - r^- = C^t - \delta c^t$$

$$b_2 \triangleq r^+ - r^- = (1 - \delta)(b + h).$$

In particular, both r^n and r^+ as defined above are greater than r^- (i.e., $b_1, b_2 > 0$) if and only if $\delta < 1$.

Proposition 9 shows that aligning the physician's incentives with those of the social planner is possible when δ is not too large and entails providing two bonuses. We start with a baseline payment level (r^-) that applies when a diagnostic test returns a negative result. The physician receives a bonus ($b_1 = C^t - \delta c^t$) over that baseline payment when the physician chooses not to order a test; the physician receives a different bonus ($b_2 = (1 - \delta)(b + h)$) over the baseline payment when the physician orders a test that returns a positive result.

An important implication from Proposition 9 is that the physician should be rewarded for not only skipping testing but also for ordering a diagnostic test that confirms a severe condition. On the one hand, under this payment system, the physician will always receive a bonus for not ordering the test. By contrast, if the physician chooses to order the test, whether she receives a bonus depends on the test's outcome. By selecting the optimal bonuses, the optimal payment scheme induces the physician to make the diagnostic effort and testing decisions that maximize social welfare: the first bonus reflects the cost saving from not ordering the test, whereas the second reflects the benefit from ordering a test that returns a positive result.

In the next corollary, we compare r^+ and r^n (i.e., b_1 and b_2).

Corollary 4. *Suppose $\delta < 1$. At the socially optimal payment scheme, $b_1 < b_2$ if and only if $\delta < (b + h - C^t)/(b + h - c^t)$ (< 1).*

Corollary 4 gives a condition that ensures the bonus for a positive test is higher than the bonus for not testing. In particular, when δ is low enough, under the fee-for-service payment system, the physician tends to order tests that do not necessarily justify the cost-benefit tradeoff. Consequently, in the redesigned payment system, to align with the social optimum, the bonus for a positive test should be set higher than the bonus for not testing, providing an incentive for the physician to order a test only when it gives "bang for the buck." When δ is too high, under the fee-for-service payment system, the physician already has an incentive to order tests. To attain the social optimum, the payer would need to provide stronger incentives for not testing than for obtaining a positive test, reflecting the focus on curbing unnecessary testing.

7. Conclusions

So far, most of the research on physician payment schemes has focused on their effect on physicians' treatment decisions. Little attention has been paid to *diagnostic* decisions (Berenson and Singh 2018). The impact is not immediately clear, especially when physicians are expected to perform a confirmatory test before diagnosing the severe condition, meaning physicians may be able to use testing to substitute for their diagnostic effort. In this paper, we develop a parsimonious model to analyze the impact of a physician payment scheme on a physician's effort and testing decisions made during the diagnostic process and thus on diagnostic accuracy and social welfare. Our paper represents an initial attempt to understand the impact of the physician payment scheme on the diagnostic decision-making process. By modeling the decisions pertaining to both diagnostic effort and confirmatory testing, our paper generates novel insights into how to design physician payment schemes in view of the intricacies of diagnostic decision-making. When a physician receives additional compensation for confirmatory testing, yet the diagnostic effort incurs a cost that is borne solely by the physician, the physician's incentives may be misaligned with those of the patient and payer.

One might expect the physician to use confirmatory testing to substitute for costly diagnostic effort (Sirovich 2011). Following this logic, increased revenue from testing would incentivize the physician to increase the use of testing and reduce the level of effort. Interestingly, we show that, depending on the additional revenue from testing, the physician can use testing and effort in either a complementary or substitutive manner. In fact, scenarios exist in which higher testing revenue motivates the physician to choose *higher* diagnostic effort. Thus, when designing the physician payment scheme, the payer must consider the potential impact on diagnostic effort.

Our analysis also demonstrates a fee-for-service payment scheme can result in higher diagnostic accuracy and diagnostic effort than at the social optimum. In terms of designing a fee-for-service payment scheme, one can use a flat payment in which the physician receives no additional payments for testing, or one can charge an additional fee for testing. In many cases, a flat compensation scheme maximizes social welfare, but not in all cases. When a physician is sufficiently patient centered, providing additional fees for testing may increase social welfare.

We also demonstrate how the practice environment can defy certain monotone effects. Although improving patient-centeredness may seem intuitive, we show doing so does not always improve accuracy, effort, or social welfare. In other words, the complex interaction between a physician's diagnostic effort and testing decisions leads to non-monotonicity in this incentive environment.

Finally, we propose an alternative payment scheme in which compensation is tied to the result of the test. We show the payment scheme can, in some cases, align decisions with the social optimum. This result sheds light on a multi-stakeholder perspective on designing proper incentives for better diagnostic services that balance individual health benefits and social welfare. Our findings have implications for a practice environment where AI and other cutting-edge technologies are used more frequently and are becoming part of the standard of care (Price et al. 2019): This alternative payment system is very much in the spirit of how providers are compensated when they use AI tools in their care delivery process (Parikh and Helmchen 2022).

Our model has some limitations. First, in line with the literature, we assume the confirmatory test is perfect. Relaxing this assumption complicates the analysis of the diagnostic outcome after the test—which is beyond the scope of our paper—but may not directionally change the key tradeoffs in the physician's diagnostic decision-making. Second, in our model, the patient's prior is the source of patient heterogeneity, and all patients bear the same cost share for the test. In practice, all patients may not have the same co-payment or co-insurance for a given test, and heterogeneity may derive from other sources, such as the benefit (harm) from a correct (incorrect) diagnosis. Third, consistent with the literature (e.g., Bester and Dahm 2017; Jelovac 2001), our model assumes that the basic exam provides an unbiased prior. However, if the basic exam is biased, it could impact the value of the confirmatory test in different ways depending on the direction and degree of the bias, as well as the relative costs of false-positive versus false-negative diagnoses. Furthermore, we assume all physicians are characterized by the same degree of patient-centeredness δ, whereas physicians may not actually all grant the same weight to the patient's benefit. Relaxing such assumptions to obtain a deeper understanding of the effect of the payment scheme on a physician's diagnostic decisions represents an interesting direction for future research.

Acknowledgments

The authors are grateful to the Department Editor, Prof. Jayashankar (Jay) Swaminathan, the Associate Editor, and two anonymous referees for their invaluable suggestions and insightful guidance throughout the constructive review process. They also benefited from comments and suggestions from Kristin Bibee, Chao-Wei Hwang, Nicos Savva, Jeffrey Scott, and Zhaowei Shi, as well as seminar participants at the London

Business School and session attendees at the 2021 MSOM Annual Meeting and the 2021 and 2022 INFORMS Annual Meetings. Attendees at the 2019 UNC Kenan-Flagler Business School Healthcare Workshop (organized by Lauren Xiaoyuan Lu, Sandeep Rath, and Brad Staats) provided insightful comments that inspired the research. Tinglong Dai was supported in part by the Johns Hopkins Discovery Award (2021–2023).

References

Abaluck J, Agha L, Kabrhel C, Raja A, Venkatesh A (2016) The determinants of productivity in medical testing: Intensity and allocation of care. *Amer. Econom. Rev.* 106(12):3730–3764.

Adida E, Bravo F (2019) Contracts for healthcare referral services: Coordination via outcome-based penalty contracts. *Management Sci.* 65(3):1322–1341.

Adida E, Mamani H, Nassiri S (2017) Bundled payment vs. fee-for-service: Impact of payment scheme on performance. *Management Sci.* 63(5):1606–1624.

Afendulis CC, Kessler DP (2007) Tradeoffs from integrating diagnosis and treatment in markets for health care. *Amer. Econom. Rev.* 97(3):1013–1020.

Aires DJ, Wick J, Shaath TS, Rajpara AN, Patel V, Badawi AH, Li C, Fraga GR, Doolittle G, Liu DY (2016) Economic costs avoided by diagnosing melanoma six months earlier justify >100 benign biopsies. *J. Drugs Dermatology* 15(5):527–532.

Alhassan S, Sayf A, Arsene C, Krayem H (2016) Suboptimal implementation of diagnostic algorithms and overuse of computed tomography-pulmonary angiography in patients with suspected pulmonary embolism. *Ann. Thoracic Medicine* 11(4):254.

Alizamir S, de Véricourt F, Sun P (2012) Diagnostic accuracy under congestion. *Management Sci.* 59(1):157–171.

Allard M, Jelovac I, Léger PT (2014) Payment mechanism and GP self-selection: Capitation versus fee for service. *Internat. J. Health Care Finance Econom.* 14(2):143–160.

Andritsos DA, Tang CS (2018) Incentive programs for reducing readmissions when patient care is co-produced. *Production Oper. Management* 27(6):999–1020.

Ayer T, Alagoz O, Stout NK (2012) OR forum—a POMDP approach to personalize mammography screening decisions. *Oper. Res.* 60(5):1019–1034.

Bastani H, Bayati M, Braverman M, Gummadi R, Johari R (2017) Analysis of Medicare pay-for-performance contracts. Technical report, EC Mechanism Design for Social Good Workshop, URL http://dx.doi.org/10.2139/ssrn.2839143.

Berenson R, Singh H (2018) Payment innovations to improve diagnostic accuracy and reduce diagnostic error. *Health Affairs* 37(11):1828–1835.

Bergeson SC, Dean JD (2006) A systems approach to patient-centered care. *JAMA* 296(23):2848.

Bertakis KD, Azari R (2011) Patient-centered care is associated with decreased health care utilization. *J. Amer. Board Family Medicine* 24(3):229–239.

Bester H, Dahm M (2017) Credence goods, costly diagnosis and subjective evaluation. *Econom. J.* 128(611):1367–1394.

Betcheva L, Erhun F, Jiang H (2021) OM forum—Supply chain thinking in healthcare: Lessons and outlooks. *Manufacturing Service Oper. Management* 23(6):1333–1353.

Bish E, Bish D, El Hajj H (2022) Disease bundling or specimen bundling? Cost- and capacity-efficient strategies for multi-disease testing with genetic assays. Working paper.

Carr T (2021) Are we screening too much for skin cancer? It's complicated. Online, https://undark.org/2021/01/25/debate-over-melanoma-diagnosis/ Accessed October 1, 2022.

Clemens J, Gottlieb JD (2014) Do physicians' financial incentives affect medical treatment and patient health? *Amer. Econom. Rev.* 104(4):1320–49.

Dai T, Akan M, Tayur S (2017) Imaging room and beyond: The underlying economics behind physicians' test-ordering behavior in outpatient services. *Manufacturing Service Oper. Management* 19(1):99–113.

Dai T, Singh S (2020) Conspicuous by its absence: Diagnostic expert testing under uncertainty. *Marketing Sci.* 39(3):540–563.

Dai T, Singh S (2022) Artificial intelligence on call: The physician's decision of whether to use AI in clinical practice. Working paper.

Dai T, Tayur S (2020) OM Forum—Healthcare operations management: A snapshot of emerging research. *Manufacturing Service Oper. Management* 22(5):869–887.

Dai T, Tayur S (2022) Designing AI-augmented healthcare delivery systems for physician buy-in and patient acceptance. *Production Oper. Management* 31(12):4443–4451.

de Verícourt F, Gurkan H (2023) Is your machine better than you? You may never know. *Management Sci.* In press.

Debo LG, Toktay LB, Van Wassenhove LN (2008) Queuing for expert services. *Management Sci.* 54(8):1497–1512.

Dugdale DC, Epstein R, Pantilat SZ (1999) Time and the patient-physician relationship. *J. General Internal Medicine* 14(S1):S34–S40.

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

Durbin E, Iyer G (2009) Corruptible advice. *Amer. Econom. J. Microeconom.* 1(2):220–42.

Eggleston K (2005) Multitasking and mixed systems for provider payment. *J. Health Econom.* 24(1):211–223.

Ellis RP, McGuire TG (1986) Provider behavior under prospective reimbursement: Cost sharing and supply. *J. Health Econom.* 5(2):129–151.

Epstein AJ, Johnson SJ (2012) Physician response to financial incentives when choosing drugs to treat breast cancer. *Internat. J. Health Care Finance Econom.* 12(4):285–302.

Fuloria PC, Zenios SA (2001) Outcomes-adjusted reimbursement in a health-care delivery system. *Management Sci.* 47(6):735–751.

Godager G, Wiesen D (2013) Profit or patients' health benefit? Exploring the heterogeneity in physician altruism. *J. Health Econom.* 32(6):1105–1116.

Goldsmith SM (2013) Cost analysis suggests overemphasis on biopsy rate for melanoma diagnosis. *J. Amer. Acad. Dermatology* 68(3):517–519.

Gromb D, Martimort D (2007) Collusion and the organization of delegated expertise. *J. Econom. Theory* 137(1):271–299.

Guo P, Tang CS, Wang Y, Zhao M (2019) The impact of reimbursement policy on social welfare, revisit rate, and waiting time in a public healthcare system: Fee-for-service versus bundled payment. *Manufacturing Service Oper. Management* 21(1):154–170.

Hajjar A, Alagoz O (2022) Personalized disease screening decisions considering a chronic condition. *Management Sci.* .

Hamid RN, McGregor SP, Siegel DM, Feldman SR (2019) Assessment of provider utilization through skin biopsy rates. *Dermatologic Surgery* 45(8):1035–1041.

Hollmann P, Jagmin C, Levy B (2021) Evaluation and management (E/M) office visits—2021. Working report, American Medical Association, URL https://www.ama-assn.org/system/files/2020-04/e-m-office-visit-changes.pdf.

Jelovac I (2001) Physicians' payment contracts, treatment decisions and diagnosis accuracy. *Health Econom.* 10(1):9–25.

Jiang H, Pang Z, Savin S (2012) Performance-based contracts for outpatient medical services. *Manufacturing Service Oper. Management* 14(4):654–669.

Jiang H, Pang Z, Savin S (2020) Performance incentives and competition in health care markets. *Production Oper. Management* 29(5):1145–1164.

Keskinocak P, Savva N (2020) A review of the healthcare-management (modeling) literature published in Manufacturing & Service Operations Management. *Manufacturing Service Oper. Management* 22(1):59–72.

Kline JA, Garrett JS, Sarmiento EJ, Strachan CC, Courtney DM (2020) Over-testing for suspected pulmonary embolism in American emergency departments. *Circulation: Cardiovascular Quality and Outcomes* 13(1).

Lee DK, Zenios SA (2012) An evidence-based incentive system for Medicare's End-Stage Renal Disease program. *Management Sci.* 58(6):1092–1105.

Lien HM, Ma CTA, McGuire TG (2004) Provider–client interactions and quantity of health care use. *J. Health Econom.* 23(6):1261–1283.

Maillart LM, Ivy JS, Ransom S, Diehl K (2008) Assessing dynamic breast cancer screening policies. *Oper. Res.* 56(6):1411–1427.

Marinoso BG, Jelovac I (2003) GPs' payment contracts and their referral practice. *J. Health Econom.* 22(4):617–635.

Mullainathan S, Obermeyer Z (2021) Diagnosing physician error: A machine learning approach to low-value health care. *Quart. J. Econom.* 137(2):679–727.

Nassiri S, Adida E, Mamani H (2022) Reference pricing for healthcare services. *Manufacturing Service Oper. Management* 24(2):921–937.

Natarajan KV, Swaminathan JM (2018) Global health. Dai T, Tayur S, eds., *Handbook of Healthcare Analytics: Theoretical Minimum for Conducting 21st Century Research on Healthcare Operations*, 137–158 (Hoboken, NJ: John Wiley & Sons, Inc.).

National Academies (2015) The path to improve diagnosis and reduce diagnostic error. Balogh EP, Miller BT, Ball JR, eds., *Improving Diagnosis in Health Care*, chapter 9 (National Academies Press).

Newhouse JP (2004) *Pricing the Priceless: A Health Care Conundrum* (Cambridge, M.A.: MIT Press).

Newman-Toker DE, McDonald KM, Meltzer DO (2013) How much diagnostic safety can we afford, and how should we decide? A health economics perspective. *BMJ Quality Safety* 22 Suppl 2:ii11–ii20.

Orfanoudaki A, Saghafian S, Song K, Chakkera HA, Cook C (2022) Algorithm, human, or the centaur: How to enhance clinical care? Working paper.

Paç MF, Veeraraghavan S (2015) False diagnosis and overtreatment in services. Working paper, University of Pennsylvania, The Wharton School.

Parikh RB, Helmchen LA (2022) Paying for artificial intelligence in medicine. *npj Digital Medicine* 5(1):63.

Price WN, Gerke S, Cohen IG (2019) Potential liability for physicians using artificial intelligence. *JAMA* 322(18):1765.

Shute N (2011) MRIs more likely to be negative when the doctor profits. *NPR*, URL https://n.pr/3KRBfSo.

Singh H, Graber ML, Hofer TP (2019) Measures to improve diagnostic safety in clinical practice. *J. Patient Safety* 15(4):311–316.

Singh H, Meyer AND, Thomas EJ (2014) The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. *BMJ Quality Safety* 23(9):727–731.

Sirovich BE (2011) Too little? Too much? Primary care physicians' views on US health care. *Arch. Internal Medicine* 171(17):1582.

Skaggs R, Coldiron B (2021) Skin biopsy and skin cancer treatment use in the medicare population, 1993 to 2016. *J. Amer. Academ. Dermatology* 84(1):53–59.

Smith JE, Ulu C (2017) Risk aversion, information acquisition, and technology adoption. *Oper. Res.* 65(4):1011–1028.

Swetter S, Geller A (2022) Melanoma: Clinical features and diagnosis. URL https://www.uptodate.com/contents/melanoma-clinical-features-and-diagnosis.

Topol E (2019) *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again* (New York: Basic Books).

Trzeciak S, Mazzarelli A (2019) *Compassionomics: The Revolutionary Scientific Evidence That Caring Makes a Difference* (Pensacola, FL: Studer Group).

US National Cancer Institute (2022) Cancer stat facts: Melanoma of the skin. Online, https://seer.cancer.gov/statfacts/html/melan.html Accessed October 1, 2022.

Vlachy J, Ayer T, Ayvaci M, Raghunathan S (2023) The business of healthcare: The role of physician integration in bundled payments. *Manufacturing Service Oper. Management* (in press).

Wahie S, Lawrence CM (2007) Wound complications following diagnostic skin biopsies in dermatology inpatients. *Archives of Dermatology* 143(10).

Welch HG, Mazer BL, Adamson AS (2021) The rapid rise in cutaneous melanoma diagnoses. *New England Journal of Medicine* 384(1):72–79.

Zhang DJ, Gurvich I, Van Mieghem JA, Park E, Young RS, Williams MV (2016) Hospital readmissions reduction program: An economic and operational analysis. *Management Sci.* 62(11):3351–3371.

Zorc S, Chick SE, Hasija S (2017) Outcomes-based reimbursement policies for chronic care pathways. Working paper, University of Virginia, Darden School of Business.

Online Appendix to "Impact of Physician Payment Scheme on Diagnostic Effort and Testing"

A1: Notations

$s \in \{\bar{s}, \underline{s}\}$	the patient's true state: severe or mild
p	the prior probability that the patient's state is severe
$f(\cdot)$, $F(\cdot)$, and μ	the pdf, cdf, and mean of the distribution of the patient priors
$e \in \{L, H\}$	the physician's effort level: low or high
c^e	the cost of effort when effort level is $e = H$
$\sigma \in \{\bar{\sigma}, \underline{\sigma}\}$	the signal (indicative or not indicative), which imperfectly reflects patient state
θ	the probability that the signal matches the true state
$t \in \{0, 1\}$	testing decision: do not test or test
c^t	the patient's cost share of the test
C^t	total cost of the test (payer and patient shares combined)
h	the penalty incurred to patient in the case of false negative diagnosis
δ	the physician's degree of patient-centeredness
b	the benefit incurred to the patient in the case of a true positive diagnosis
r^t, r^n	physician compensation, based on testing decision: with test or without test
Δr	the net payment differential: $r^t - r^n$
U_{patient}	the patient's utility
$\Pi_{\text{physician}}$	the physician's payoff
$U_{\text{physician}}$	the physician's utility
Π_{payer}	the payer's payoff
SW	social welfare
$\bar{c}^F, \bar{c}^S, \bar{c}^D$	thresholds on the cost of effort
$p_1^F, p_2^F, p_1^S, p_2^S$, p_1^D, p_2^D	thresholds on patient prior probability
r^+, r^-	parameters of the diagnosis-based payment scheme

Table A1 Notations

A2: Proofs of Technical Results

Proof of Lemma 1. The physician should order a test if $r^t + \delta(pb - c^t) > r^n - \delta ph$, which is equivalent to

$$p > \frac{\delta c^t - \Delta r}{\delta(b+h)},$$

and provide a diagnosis of mild otherwise.

We have $0 < \frac{\delta c^t - \Delta r}{\delta(b+h)} < 1$ if and only if

$$\delta(c^t - b - h) < \Delta r < \delta c^t. \tag{A1}$$

If $\Delta r \leq \delta(c^t - b - h)$, the physician never orders a test, regardless of p. If $\Delta r \geq \delta c^t$, the physician orders a test for all patients. *Q.E.D.*

Proof of Lemma 2. If the consultation generates a positive signal, the physician updates the patient's probability of suffering from the condition:

$$\Pr(s = \bar{s} | \sigma = \bar{\sigma}) = \frac{\theta p}{\theta p + (1-\theta)(1-p)}.$$

The physician should order a test if

$$r^t + \delta \cdot \frac{\theta p}{\theta p + (1-\theta)(1-p)} \cdot b - \delta c^t > r^n - \delta \cdot \frac{\theta p}{\theta p + (1-\theta)(1-p)} \cdot h,$$

which gives

$$p > \frac{(1-\theta)(\delta c^t - \Delta r)}{(1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t) + \Delta r]}.$$

We have $0 < \frac{(1-\theta)(\delta c^t - \Delta r)}{(1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t)+\Delta r]} < 1$ if and only if $\delta(c^t - b - h) < \Delta r < \delta c^t$, which coincides with (A1). If $\Delta r \leq \delta(c^t - b - h)$, the physician never orders a test, regardless of p. If $\Delta r \geq \delta c^t$, the physician orders a test for all patients.

If the consultation generates a negative signal, the physician updates the patient's probability of suffering from the condition:

$$\Pr(s = \bar{s} | \sigma = \underline{\sigma}) = \frac{(1-\theta)p}{(1-\theta)p + \theta(1-p)}.$$

The physician should order a test if

$$r^t + \delta \cdot \frac{(1-\theta)p}{(1-\theta)p + \theta(1-p)} \cdot b - \delta c^t > r^n - \delta \cdot \frac{(1-\theta)p}{(1-\theta)p + \theta(1-p)} \cdot h,$$

which is equivalent to

$$p > \frac{\theta(\delta c^t - \Delta r)}{\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t) + \Delta r]}.$$

We have $0 < \frac{\theta(\delta c^t - \Delta r)}{\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t)+\Delta r]} < 1$ if and only if $\delta(c^t - b - h) < \Delta r < \delta c^t$, which coincides with (A1). If $\Delta r \leq \delta(c^t - b - h)$, the physician never orders a test, regardless of p. If $\Delta r \geq \delta c^t$, the physician orders a test for all patients.

If $\delta(c^t - b - h) < \Delta r < \delta c^t$, the inequality

$$\frac{(1-\theta)(\delta c^t - \Delta r)}{(1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t) + \Delta r]} < \frac{\delta c^t - \Delta r}{\delta(b+h)} < \frac{\theta(\delta c^t - \Delta r)}{\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t) + \Delta r]}$$

can be obtained from simple algebra using the condition $\theta > 1/2$. *Q.E.D.*

PROOF OF LEMMA 3. We consider the following cases:

(i) If $\Delta r \leq \delta(c^t - b - h)$ or $\Delta r \geq \delta c^t$, based on Lemmas 1 and 2, the final decision is the same under low and high effort (regardless of the signal), so the physician has no incentive to exert costly high effort.

(ii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$, the physician's effort decision can be determined as follows, according to Lemmas 1 and 2:

(a) If $\frac{\theta(\delta c^t - \Delta r)}{\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t)+\Delta r]} < p < 1$, the physician orders the test regardless of the effort or the signal (if applicable). Thus, the physician chooses not to exert high effort.

(b) If $0 < p \leq \frac{(1-\theta)(\delta c^t - \Delta r)}{(1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t)+\Delta r]}$, the physician provides a mild diagnosis regardless of the effort or the signal. Thus, the physician chooses not to exert high effort.

(c) If $\frac{(1-\theta)(\delta c^t - \Delta r)}{(1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t)+\Delta r]} < p \leq \frac{\theta(\delta c^t - \Delta r)}{\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t)+\Delta r]}$, the physician follows the signal if exerting high effort. The condition for the physician to exert high effort is as follows:

i. If $\frac{(1-\theta)(\delta c^t-\Delta r)}{(1-\theta)(\delta c^t-\Delta r)+\theta[\delta(b+h-c^t)+\Delta r]} < p \leq \frac{\delta c^t-\Delta r}{\delta(b+h)}$, the physician exerts high effort if and only if

$$[\theta p+(1-\theta)(1-p)]\cdot\left[r^t+\delta\cdot\frac{\theta p}{\theta p+(1-\theta)(1-p)}\cdot b-\delta c^t\right]$$
$$+[(1-\theta)p+\theta(1-p)]\cdot\left[r^n-\delta\cdot\frac{(1-\theta)p}{(1-\theta)p+\theta(1-p)]}\cdot h\right]-c^e>r^n-\delta ph,$$

which gives

$$p>\frac{(1-\theta)(\delta c^t-\Delta r)+c^e}{(1-\theta)(\delta c^t-\Delta r)+\theta[\delta(b+h-c^t)+\Delta r]}=p_1^F.$$

We can easily verify that

$$p_1^F>\frac{(1-\theta)(\delta c^t-\Delta r)}{(1-\theta)(\delta c^t-\Delta r)+\theta[\delta(b+h-c^t)+\Delta r]}.$$

Moreover, we have

$$p_1^F\leq\frac{\delta c^t-\Delta r}{\delta(b+h)},$$

if and only if

$$c^e\leq(2\theta-1)(\delta c^t-\Delta r)\left[1-\frac{\delta c^t-\Delta r}{\delta(b+h)}\right]=\bar{c}^F. \tag{A2}$$

If c^e is above the threshold $\bar{c}^F$, the physician does not exert high effort for any patient within this range of priors.

ii. If $\frac{\delta c^t-\Delta r}{\delta(b+h)} < p \leq \frac{\theta(\delta c^t-\Delta r)}{\theta(\delta c^t-\Delta r)+(1-\theta)[\delta(b+h-c^t)+\Delta r]}$, the physician exerts high effort if and only if

$$[\theta p+(1-\theta)(1-p)]\cdot\left[r^t+\delta\cdot\frac{\theta p}{\theta p+(1-\theta)(1-p)}\cdot b-\delta c^t\right]$$
$$+[(1-\theta)p+\theta(1-p)]\cdot\left[r^n-\delta\cdot\frac{(1-\theta)p}{(1-\theta)p+\theta(1-p)]}\cdot h\right]-c^e$$
$$>r^t+\delta pb-\delta c^t,$$

which gives

$$p<\frac{\theta(\delta c^t-\Delta r)-c^e}{\theta(\delta c^t-\Delta r)+(1-\theta)[\delta(b+h-c^t)+\Delta r]}=p_2^F.$$

We can easily verify that

$$p_2^F<\frac{\theta(\delta c^t-\Delta r)}{\theta(\delta c^t-\Delta r)+(1-\theta)[\delta(b+h-c^t)+\Delta r]}.$$

Moreover, we have

$$p_2^F\geq\frac{\delta c^t-\Delta r}{\delta(b+h)},$$

if and only if

$$c^e\leq(2\theta-1)(\delta c^t-\Delta r)\left[1-\frac{\delta c^t-\Delta r}{\delta(b+h)}\right]=\bar{c}^F,$$

which coincides with (A2). If c^e is above the threshold, the physician exerts low effort for any patient within this range of priors. *Q.E.D.*

Proof of Proposition 3 We denote by $a^F(p)$ (resp. $a^S(p)$) the diagnostic accuracy for a patient with a prior p under fee-for-service (resp. in the social optimum). To prove the result, we proceed by stating and proving three intermediary results. We first establish the following corollary:

Corollary A1. *Under fee-for-service, the diagnostic accuracy is as follows:*

(i) If $\Delta r \le \delta(c^t - b - h)$*, then* $a^F(p) = 1 - p$*;*

(ii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ *and* $c^e > \bar{c}^F$*, then*

$$a^F(p) = \begin{cases} 1-p & \textit{if } p < \frac{\delta c^t - \Delta r}{\delta(b+h)} \\ 1 & \textit{else;} \end{cases}$$

(iii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ *and* $c^e \le \bar{c}^F$*, then*

$$a^F(p) = \begin{cases} 1-p & \textit{if } p \le p_1^F \\ 1-p(1-\theta) & \textit{if } p_1^F < p \le p_2^F \\ 1 & \textit{if } p > p_2^F; \end{cases}$$

(iv) If $\Delta r \ge \delta c^t$*, then* $a^F(p) = 1$*.*

Proof of Corollary A1. The four parts of the corollary correspond to those in Proposition 1. Note from Proposition 1 that in case (i), the physician always reaches a mild diagnosis after exerting low effort. Thus, the diagnostic accuracy corresponds to the likelihood that the patient's true state is negative; that is, $1-p$. Similarly, in case (iv), the physician always orders the test. Because the test can perfectly reveal the patient's true state, the diagnostic accuracy is 1.

In case (ii), the physician exerts low effort for all patients and orders a test if and only if $p \ge \frac{\delta c^t - \Delta r}{\delta(b+h)}$. Thus, if $p \ge \frac{\delta c^t - \Delta r}{\delta(b+h)}$, the diagnostic accuracy is 1; otherwise, the physician reaches a mild diagnosis, so the diagnostic accuracy is the likelihood that the patient's true state is negative, that is, $1-p$.

In case (iii), if $p \le p_1^F$, the physician exerts low effort and reaches a mild diagnosis; thus, the diagnostic accuracy is $1-p$. Next, if $p_1^F < p \le p_2^F$, the physician exerts high effort and follows the signal obtained; thus, the diagnostic accuracy is

$$\begin{aligned}
&\underbrace{[p\theta + (1-p)(1-\theta)]}_{\Pr(\sigma=\bar{\sigma})} \cdot 1 + \underbrace{[p(1-\theta) + (1-p)\theta]}_{\Pr(\sigma=\underline{\sigma})} \cdot \underbrace{\frac{(1-p)\theta}{p(1-\theta)+p(1-\theta)}}_{\Pr(s=\underline{s}|\sigma=\underline{\sigma})} \\
&= p\theta + (1-p)(1-\theta) + (1-p)\theta \\
&= 1 - p(1-\theta).
\end{aligned}$$

Finally, if $p > p_2^F$, the physician orders a test, so the diagnostic accuracy is 1. *Q.E.D.*

We next establish the following result, where we define $Q(x) = \int_0^x t f(t) dt$ and $\mu = Q(1)$.

Proposition A1. *Under FFS, the average population diagnostic accuracy is as follows:*

(i) If $\Delta r \le \delta(c^t - b - h)$*, then* $\mathbb{E}_p[a^F(p)] = 1 - \mu$*;*

(ii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ *and* $c^e > \bar{c}^F$*, then*

$$\mathbb{E}_p[a^F(p)] = 1 - Q\left(\frac{\delta c^t - \Delta r}{\delta(b+h)}\right);$$

(iii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ *and* $c^e \le \bar{c}^F$*, then*

$$\mathbb{E}_p[a^F(p)] = 1 - \theta Q(p_1^F) - (1-\theta) Q(p_2^F);$$

(iv) If $\Delta r \geq \delta c^t$, then $\mathbb{E}_p[a^F(p)] = 1$.

In addition, the average population diagnostic accuracy is continuous and non-decreasing in Δr.

Proof of Proposition A1 The expression of $\mathbb{E}_p[a^F(p)]$ follows directly from Corollary A1. It remains to show the monotonicity. For $\Delta r \leq \delta(c^t - b - h)$, $\mathbb{E}_p[a^F(p)]$ is a constant (case (i)). For $\delta(c^t - b - h) < \Delta r < \delta c^t$, at the left extreme of the range, $\delta(c^t - b - h)^+$, we have $\bar{c}^F = 0^+$, and thus $c^e > \bar{c}^F$, so we are in case (ii). Moreover, at this extreme, $\mathbb{E}_p[a^F(p)] = (1-\mu)^+$, so the expected accuracy is continuous. Because $\Delta r < \delta c^t$ in this range,

$$\frac{\partial \mathbb{E}_p[a^F(p)]}{\partial \Delta r} = \frac{\delta c^t - \Delta r}{\delta^2 (b+h)^2} f\left(\frac{\delta c^t - \Delta r}{\delta(b+h)}\right) > 0,$$

so the expected accuracy is increasing while in case (ii). If $c^e > (2\theta - 1)\delta(b+h)/4$ (which is the maximum value taken by $\bar{c}^F$), we remain in case (ii) over the whole range, and at $(\delta c^t)^-$, $\mathbb{E}_p[a^F(p)] = 1^-$, so the aggregate accuracy is continuous and non-decreasing. Otherwise, we switch to case (iii) within the range $\delta(c^t - b - h) < \Delta r < \delta c^t$, and then back to case (ii) because $\bar{c}^F$ is unimodal. The switch occurs when Δr is such that $c^e = \bar{c}^F$. When $c^e = \bar{c}^F$, we have (see proof of Lemma 3) $p_1^F = p_2^F = (\delta c^t - \Delta r)/(\delta(b+h))$; thus, the aggregate accuracy is continuous. Monotonicity in case (iii) follows from p_1^F and p_2^F decreasing in Δr (see proof of Corollary 1). *Q.E.D.*

We next establish the following result:

Proposition A2. *In the social optimum, the average population diagnostic accuracy is as follows:*

(i) If $b + h \leq C^t$, then $\mathbb{E}_p[a^S(p)] = 1 - \mu$;

(ii) If $b + h > C^t$ and $c^e > \bar{c}^S$, then

$$\mathbb{E}_p[a^S(p)] = 1 - Q\left(\frac{C^t}{b+h}\right) \quad (> 1 - \mu);$$

(iii) If $b + h > C^t$ and $c^e \leq \bar{c}^S$, then

$$\mathbb{E}_p[a^S(p)] = 1 - \theta Q(p_1^S) - (1-\theta) Q(p_2^S) \quad \left(\geq 1 - Q\left(\frac{C^t}{b+h}\right) > 1 - \mu\right).$$

Proof of Proposition A2. The result follows similarly to Proposition A1 after noting the socially optimal policy is as follows:

(i) If $b + h \leq C^t$, then $a^S(p) = 1 - p$.

(ii) If $b + h > C^t$ and $c^e > \bar{c}^S$, then

$$a^S(p) = \begin{cases} 1 - p & \text{if } p < \frac{C^t}{b+h} \\ 1 & \text{else.} \end{cases}$$

(iii) If $b + h > C^t$ and $c^e \leq \bar{c}^S$, then

$$a^S(p) = \begin{cases} 1 - p & \text{if } p \leq p_1^S \\ 1 - p(1-\theta) & \text{if } p_1^S < p \leq p_2^S \\ 1 & \text{if } p > p_2^S. \end{cases}$$

Q.E.D.

Proposition 3 then follows from the above intermediary results. *Q.E.D.*

Proof of Corollary 1. We first establish a preliminary result: if $\delta(c^t - b - h) < \Delta r < \delta c^t$, then p_1^F and p_2^F are decreasing in Δr and $\bar{c}^F$ is unimodal in Δr, reaching a maximum equal to $(2\theta-1)\delta(b+h)/4$ at $\Delta r = \delta(c^t - (b+h)/2)$.

To show this preliminary result, note we have

$$\begin{aligned}\frac{\partial p_1^F}{\partial \Delta r} &= \frac{-(1-\theta)((1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t)+\Delta r]) - [(1-\theta)(\delta c^t - \Delta r) + c^e](2\theta-1)}{((1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t)+\Delta r])^2} \\ &= \frac{-(1-\theta)\theta\delta(b+h) - c^e(2\theta-1)}{((1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t)+\Delta r])^2} < 0,\end{aligned}$$

where the last inequality follows from $\theta > 1/2$. Similarly,

$$\begin{aligned}\frac{\partial p_2^F}{\partial \Delta r} &= \frac{-\theta(\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t)+\Delta r]) - [\theta(\delta c^t - \Delta r) - c^e](1-2\theta)}{(\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t)+\Delta r])^2} \\ &= \frac{-\theta(1-\theta)\delta(b+h) - c^e(2\theta-1)}{(\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t)+\Delta r])^2} < 0. \\ \frac{\partial \bar{c}^F}{\partial \Delta r} &= (2\theta-1)\left(-1 + \frac{\delta c^t - \Delta r}{\delta(b+h)} + (\delta c^t - \Delta r)\frac{1}{\delta(b+h)}\right) \\ &= (2\theta-1)\left(-1 + 2\frac{\delta c^t - \Delta r}{\delta(b+h)}\right).\end{aligned}$$

Therefore, $\bar{c}^F$ increases in Δr when $\Delta r < \delta(c^t - \frac{b+h}{2})$, and $\bar{c}^F$ decreases in Δr when $\Delta r > \delta(c^t - \frac{b+h}{2})$. By substitution, we find it reaches a maximum equal to $(2\theta-1)\delta(b+h)/4$ at $\Delta r = \delta(c^t - (b+h)/2)$. This concludes the proof of the preliminary result.

Hence, $c^e \le \bar{c}^F$ can occur only when c^e is below the maximum value reached by $\bar{c}^F$. In that case, $c^e \le \bar{c}^F$ occurs whenever Δr is between the two roots of the quadratic equation $c^e = \bar{c}^F$; that is, denoting $x = \delta c^t - \Delta r$,

$$\begin{aligned} c^e &= (2\theta-1)x\left(1 - \frac{x}{\delta(b+h)}\right) \\ \Leftrightarrow & x^2 - \delta(b+h)x + \frac{c^e\delta(b+h)}{2\theta-1} = 0.\end{aligned}$$

The result follows from the standard solution of a quadratic equation. *Q.E.D.*

Proof of Corollary 2. When $\delta(c^t - b - h) < \Delta r < \delta c^t$ and $c^e \le \bar{c}^F$, we show from the proof of Proposition 4 that the range of the priors that correspond to high effort is

$$\begin{aligned} p_2^F - p_1^F &= \frac{\theta(\delta c^t - \Delta r) - c^e}{\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t)+\Delta r]} - \frac{(1-\theta)(\delta c^t - \Delta r) + c^e}{(1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t)+\Delta r]} \\ &= \frac{\bar{c}^F - c^e}{\theta(1-\theta)\delta(b+h) + (2\theta-1)\bar{c}^F}\end{aligned}$$

after simplifications. Moreover,

$$\begin{aligned}\frac{\partial(p_2^F - p_1^F)}{\partial \bar{c}^F} &= \frac{\theta(1-\theta)\delta(b+h) + (2\theta-1)\bar{c}^F - (2\theta-1)(\bar{c}^F - c^e)}{(\theta(1-\theta)\delta(b+h) + (2\theta-1)\bar{c}^F)^2} \\ &= \frac{\theta(1-\theta)\delta(b+h) + (2\theta-1)c^e}{(\theta(1-\theta)\delta(b+h) + (2\theta-1)\bar{c}^F)^2} > 0.\end{aligned}$$

Thus, it suffices to use the monotonicity results on $\bar{c}^F$. From the above preliminary result, we know that $\bar{c}^F$ increases in Δr when $\Delta r < \delta(c^t - \frac{b+h}{2})$, and $\bar{c}^F$ decreases in Δr when $\Delta r > \delta(c^t - \frac{b+h}{2})$. By substitution, we obtain that for $\Delta r = \delta(c^t - \frac{b+h}{2})$, the range $p_2^F - p_1^F$ equals $2\theta - 1 - 4c^e/(\delta(b+h))$. *Q.E.D.*

Proof of Proposition 4. From the proof of Corollary 1, the widest range of high effort under the physician's optimal strategy is obtained for $\Delta r = \delta(c^t - \frac{b+h}{2})$, where $p_2^F - p_1^F = 2\theta - 1 - 4c^e/(\delta(b+h))$ and where $\bar{c}^F = (2\theta-1)\delta(b+h)/4$. Therefore, we want to determine the sign of $p_2^S - p_1^S - (2\theta - 1 - 4c^e/(\delta(b+h)))$. Using derivations from the proof of Corollary 1, we have

$$p_2^S - p_1^S = \frac{\bar{c}^S - c^e}{\theta(1-\theta)(b+h) + (2\theta-1)\bar{c}^S}.$$

As shown in the proof of Corollary 1, the expression for $p_2^S - p_1^S$ is increasing in $\bar{c}^S$. Thus, the socially optimum range of priors leading to high effort reaches a maximum (over C^t) when $\bar{c}^S$ is at its maximum, that is, for $C^t = (b+h)/2$, and $\bar{c}^S$ then equals $(2\theta-1)(b+h)/4$, and $p_2^S - p_1^S$ then equals $2\theta - 1 - 4c^e/(b+h)$.

When $\delta > 1$, the biggest FFS range of priors leading to high effort is bigger than the socially optimal largest range of priors leading to high effort; thus, it is bigger than the socially optimum range of priors leading to high effort for any C^t.

When $\delta \le 1$, the biggest FFS range of priors leading to high effort is bigger than the social optimal range of priors leading to high effort when C^t is outside the two roots of the equation $p_2^S - p_1^S - (2\theta - 1 - 4c^e/(\delta(b+h))) = 0$, which is quadratic in C^t. The result follows from the standard solution of a quadratic equation. *Q.E.D.*

Proof of Proposition 5 To prove this result, we start with showing an intermediate result, where $\bar{F}(\cdot) = 1 - F(\cdot)$ and $Q(x) = \int_0^x tf(t)dt$:

Proposition A3. *Under fee-for-service, the average population social welfare is as follows:*

(i) If $\Delta r \le \delta(c^t - b - h)$, *then* $\mathbb{E}_p[SW^F(p)] = -h\mu$;

(ii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ *and* $c^e > \bar{c}^F$, *then*

$$\mathbb{E}_p[SW^F(p)] = b\mu - C^t \bar{F}\left(\frac{\delta c^t - \Delta r}{\delta(b+h)}\right) - (b+h)Q\left(\frac{\delta c^t - \Delta r}{\delta(b+h)}\right);$$

(iii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ *and* $c^e \le \bar{c}^F$, *then*

$$\begin{aligned}\mathbb{E}_p[SW^F(p)] = {} & b\mu - C^t - c^e[F(p_2^F) - F(p_1^F)] + C^t[\theta F(p_2^F) + (1-\theta)F(p_1^F)] \\ & - C^t\left[\theta Q(p_2^F) + (1-\theta)Q(p_1^F)\right] - (b+h-C^t)\left[(1-\theta)Q(p_2^F) + \theta Q(p_1^F)\right];\end{aligned}$$

(iv) If $\Delta r \ge \delta c^t$, *then* $\mathbb{E}_p[SW^F(p)] = b\mu - C^t$.

Proof of Proposition A3 The result follows from noting that for a patient with given prior p, the social welfare is as follows:

(i) If $\Delta r \le \delta(c^t - b - h)$, then $SW^F(p) = -ph$;

(ii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ and $c^e > \bar{c}^F$, then

$$SW^F(p) = \begin{cases} -ph & \text{if } p < \frac{\delta c^t - \Delta r}{\delta(b+h)} \\ pb - C^t & \text{else;} \end{cases}$$

(iii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ and $c^e \le \bar{c}^F$, then

$$SW^F(p) = \begin{cases} -ph & \text{if } p \le p_1^F \\ -c^e - C^t(p\theta + (1-p)(1-\theta)) + bp\theta - hp(1-\theta) & \text{if } p_1^F < p \le p_2^F \\ pb - C^t & \text{if } p > p_2^F; \end{cases},$$

(iv) If $\Delta r \geq \delta c^t$, then $SW^F(p) = pb - C^t$. *Q.E.D.*

We now return to the proof of Proposition 5. If $c^e > (2\theta - 1)\delta(b+h)/4$, case (iii) does not occur. As Δr increases from $-\infty$ to $+\infty$, we go from case (i) to case (ii) to case (iv) of Proposition A3. Taking the derivative of $\mathbb{E}_p[SW^F(p)]$ w.r.t. Δr in case (ii), we obtain

$$\frac{-\Delta r - \delta(C^t - c^t)}{\delta^2(b+h)} f\left(\frac{\delta c^t - \Delta r}{\delta(b+h)}\right).$$

It follows that $\mathbb{E}_p[SW^F(p)]$ is unimodal in case (ii), reaching a maximum for $\Delta r = -\delta(C^t - c^t)$. This value is in the range of case (ii) if $b + h > C^t$. (Otherwise, social welfare is decreasing over case (ii).) By simple substitution, we find the value reached by $\mathbb{E}_p[SW^F(p)]$ at $\Delta r = -\delta(C^t - c^t)$ is

$$b\mu - C^t \bar{F}\left(\frac{C^t}{b+h}\right) - (b+h)Q\left(\frac{C^t}{b+h}\right).$$

We next prove this quantity is greater than both (1) $-h\mu$ (value in case (i)) and (2) $b\mu - C^t$ (value in case (iv)):

- Proof of (1): We have

$$xF(x) = \int_0^x x f(t)dt > \int_0^x t f(t)dt = Q(x).$$

The result follows after substituting x with $C^t/(b+h)$.

- Proof of (2): We have

$$Q(1) - Q(x) = \int_x^1 t f(t)dt > \int_x^1 x f(t)dt = x\bar{F}(x).$$

The result follows after substituting x with $C^t/(b+h)$.

The socially optimal social welfare can be obtained using Proposition A3 with the values $\delta = 1$ and $\Delta r = -(C^t - c^t)$. Hence, if $b + h \leq C^t$, the maximum social welfare is $-h\mu$, which matches the FFS case. If $b + h > C^t$, we are in case (ii) when $c^e > \bar{c}^S$, and the maximum social welfare matches FFS. *Q.E.D.*

Proof of Lemma 4. Note that, for $\delta c^t = \Delta r$, $\bar{c}^F = 0 < c^e$. Moreover, we obtain

$$\frac{\partial \bar{c}^F}{\partial \delta} = c^t\left(1 - \frac{c^t}{b+h}\right) + \frac{\Delta r^2}{\delta^2(b+h)} > 0;$$
$$\lim_{\delta \to \infty} \bar{c}^F = \infty.$$

Hence, when $\Delta r > 0$, it can be seen from Proposition A1 that, as δ increases from 0, we go from case (iv) (where the diagnostic accuracy is constant) to case (ii) (as δ passes the value $\Delta r/c^t$) to case (iii) (as $\bar{c}^F$ passes the value c^e). In case (ii), using the expression in Proposition A1(ii), the diagnostic accuracy is decreasing in δ, because $Q(\cdot)$ is increasing. For case (iii), we need to consider how p_1^F and p_2^F vary with δ. We find

$$\frac{\partial p_1^F}{\partial \delta} = \frac{\theta(1-\theta)\Delta r(b+h) - c^e[(1-\theta)c^t + \theta(b+h-c^t)]}{((1-\theta)(\delta c^t - \Delta r) + \theta[\delta(b+h-c^t) + \Delta r])^2},$$

which is positive under the assumption that $c^e < \theta(1-\theta)\Delta r(b+h)/((1-\theta)c^t + \theta(b+h-c^t))$. Moreover,

$$\frac{\partial p_2^F}{\partial \delta} = \frac{\theta(1-\theta)\Delta r(b+h) + c^e[\theta c^t + (1-\theta)(b+h-c^t)]}{(\theta(\delta c^t - \Delta r) + (1-\theta)[\delta(b+h-c^t) + \Delta r])^2} > 0.$$

Therefore, in the situation considered in the lemma, both p_1^F and p_2^F are increasing in δ. Using the expression in Proposition A1(iii), because $Q(\cdot)$ is increasing, it follows that in case (iii), the diagnostic accuracy is decreasing in δ. *Q.E.D.*

Proof of Lemma 5. The proof of Lemma 4 shows $\bar{c}^F$ increases from zero to infinity as δ increases. Hence, when $\Delta r = 0$, Proposition A1 shows that as δ increases from 0, we go from case (ii) to case (iii) (as $\bar{c}^F$ passes the value c^e). In case (ii), using the expression in Proposition A1(ii), the diagnostic accuracy is independent of δ. In case (iii), $Q(\cdot)$ is increasing, but we now have p_1^F decreasing in δ while p_2^F is increasing in δ. We thus obtain the derivative of the diagnostic accuracy (denoting $D_1 \triangleq (1-\theta)c^t + \theta(b+h-c^t)$ and $D_2 \triangleq \theta c^t + (1-\theta)(b+h-c^t)$):

$$\begin{aligned}\frac{\partial \mathbb{E}_p[a^F(p)]}{\partial \delta} =& -\theta p_1^F f(p_1^F)\frac{\partial p_1^F}{\partial \delta} - (1-\theta)p_2^F f(p_2^F)\frac{\partial p_2^F}{\partial \delta}\\ =&\theta \frac{c^e}{\delta^2}\frac{(1-\theta)c^t + c^e/\delta}{D_1^2} f(p_1^F) - (1-\theta)\frac{c^e}{\delta^2}\frac{\theta c^t - c^e/\delta}{D_2^2} f(p_2^F)\\ =&\frac{c^e}{\delta^2}\theta(1-\theta)c^t\left(\frac{f(p_1^F)}{D_1^2} - \frac{f(p_2^F)}{D_2^2}\right) + \frac{(c^e)^2}{\delta^3}\left(\frac{\theta f(p_1^F)}{D_1^2} + \frac{(1-\theta)f(p_2^F)}{D_2^2}\right).\end{aligned}$$

The diagnostic accuracy is increasing in case (iii) when $f(p_1^F)/D_1^2 \geq f(p_2^F)/D_2^2$.

Otherwise, the derivative of the accuracy has the sign of

$$-\theta(1-\theta)c^t\left(\frac{f(p_2^F)}{D_2^2} - \frac{f(p_1^F)}{D_1^2}\right) + \frac{c^e}{\delta}\left(\frac{\theta f(p_1^F)}{D_1^2} + \frac{(1-\theta)f(p_2^F)}{D_2^2}\right),$$

which is positive if and only if

$$\delta < \frac{c^e}{\theta(1-\theta)c^t}\left(\frac{\theta f(p_1^F)}{D_1^2} + \frac{(1-\theta)f(p_2^F)}{D_2^2}\right)\frac{1}{\frac{f(p_2^F)}{D_2^2} - \frac{f(p_1^F)}{D_1^2}}.$$

Q.E.D.

Proof of Proposition 6. To prove this result, we establish a more detailed result below.

Proposition A4. *The range of priors leading to high effort under fee-for-service is monotonically increasing in δ if and only if*

1. $\Delta r = 0$, *or*
2. $\Delta r > 0$ *and* $K \geq 0$, *or*
3. $\Delta r > 0$ *and* $K < 0$ *and*

$$-\frac{\theta(1-\theta)(b+h)}{c^e} + \frac{1}{\Delta r}\sqrt{\frac{-K(b+h)}{c^e}} \leq 0,$$

where

$$K \triangleq -\frac{\theta^2(1-\theta)^2(\Delta r)^2(b+h)}{c^e} + \frac{c^e}{b+h}(2\theta-1)^2 c^t(b+h-c^t) + c^e\theta(1-\theta)(b+h) + \theta(1-\theta)(2\theta-1)\Delta r(b+h-2c^t).$$

Otherwise, the range of priors leading to high effort is unimodal in δ, increasing if and only if

$$\delta < \frac{2\theta-1}{-\frac{\theta(1-\theta)(b+h)}{c^e} + \frac{1}{|\Delta r|}\sqrt{\frac{-K(b+h)}{c^e}}}.$$

Proof of Proposition A4. The derivative of p_1^F w.r.t. δ has the sign of

$$\theta(1-\theta)\Delta r(b+h) - c^e((1-\theta)c^t + \theta(b+h-c^t)).$$

The derivative of p_2^F w.r.t. δ has the sign of

$$\theta(1-\theta)\Delta r(b+h) + c^e(\theta c^t + (1-\theta)(b+h-c^t)).$$

The derivative of $\bar{c}^F$ w.r.t. δ has the sign of

$$\left(\frac{\Delta r}{\delta}\right)^2 + c^t(b+h-c^t) > 0.$$

From the proof of Corollary 1, we have

$$p_2^F - p_1^F = \frac{\bar{c}^F - c^e}{\theta(1-\theta)\delta(b+h) + (2\theta-1)\bar{c}^F}.$$

Thus, $\partial(p_2^F - p_1^F)/\partial\delta$ has the sign of

$$\begin{aligned}
&\frac{\partial\bar{c}^F}{\partial\delta}[\theta(1-\theta)\delta(b+h) + (2\theta-1)\bar{c}^F] - (\bar{c}^F - c^e)\left[\theta(1-\theta)(b+h) + (2\theta-1)\frac{\partial\bar{c}^F}{\partial\delta}\right]\\
=&\frac{\partial\bar{c}^F}{\partial\delta}[\theta(1-\theta)\delta(b+h) + \bar{c}^F(2\theta-1)] - (\bar{c}^F - c^e)\theta(1-\theta)(b+h)\\
=&\frac{2\theta-1}{b+h}\left[\left(\frac{\Delta r}{\delta}\right)^2 + c^t(b+h-c^t)\right][\theta(1-\theta)\delta(b+h) + \bar{c}^F(2\theta-1)] + c^e\theta(1-\theta)(b+h)\\
&-\theta(1-\theta)(2\theta-1)(b+h)\delta\left(c^t - \frac{\Delta r}{\delta}\right)\left[b+h-c^t + \frac{\Delta r}{\delta}\right]\\
=&\frac{c^e}{b+h}(2\theta-1)^2\left(\frac{\Delta r}{\delta}\right)^2 + 2(2\theta-1)\theta(1-\theta)\delta\left(\frac{\Delta r}{\delta}\right)^2 + \frac{c^e}{b+h}(2\theta-1)^2c^t(b+h-c^t) + c^e\theta(1-\theta)(b+h)\\
&+\theta(1-\theta)(2\theta-1)\Delta r(b+h-2c^t).
\end{aligned}$$

This expression is positive when

$$c^e > \frac{(b+h)\theta(1-\theta)(2\theta-1)\Delta r\left[-(b+h-2c^t) - 2\frac{\Delta r}{\delta}\right]}{\theta(1-\theta)(b+h)^2 + (2\theta-1)^2\left[\left(\frac{\Delta r}{\delta}\right)^2 + c^t(b+h-c^t)\right]}. \tag{A3}$$

The above is necessarily true when $\Delta r = 0$.

We can write differently that $\partial(p_2^F - p_1^F)/\partial\delta$ is positive when

$$\frac{c^e}{b+h}(\Delta r)^2\left(\frac{2\theta-1}{\delta} + \frac{\theta(1-\theta)(b+h)}{c^e}\right)^2 + K > 0,$$

where

$$K \triangleq -\frac{\theta^2(1-\theta)^2(\Delta r)^2(b+h)}{c^e} + \frac{c^e}{b+h}(2\theta-1)^2c^t(b+h-c^t) + c^e\theta(1-\theta)(b+h) + \theta(1-\theta)(2\theta-1)\Delta r(b+h-2c^t).$$

Hence, if $K \geq 0$, $p_2^F - p_1^F$ is increasing in δ. If $K < 0$, $p_2^F - p_1^F$ is increasing in δ if and only if:

$$-\frac{\theta(1-\theta)(b+h)}{c^e} + \frac{1}{|\Delta r|}\sqrt{\frac{-K(b+h)}{c^e}} > 0 \text{ and } \delta < \frac{2\theta-1}{-\frac{\theta(1-\theta)(b+h)}{c^e} + \frac{1}{|\Delta r|}\sqrt{\frac{-K(b+h)}{c^e}}}.$$

Q.E.D.

Proposition 6 directly follows from Proposition A4. *Q.E.D.*

Proof of Corollary 3 If $c^t \le (b+h)/2$, the right-hand-side in (A3) is negative (because the numerator is negative), and thus, the inequality (A3) holds.

Note if $c^t = 0$, from Proposition 1, regardless of δ, the physician will exert a low diagnostic effort level and order a test for all patients, so the range will remain empty. *Q.E.D.*

Proof of Lemma 6 As detailed in the proof of Lemma 5, we go from case (ii) to case (iii) as δ increases. Using Proposition A3, we observe that social welfare is independent of δ in case (ii) when $\Delta r = 0$. We now focus on case (iii). We denote $D_1 \triangleq (1-\theta)c^t + \theta(b+h-c^t)$ and $D_2 \triangleq \theta c^t + (1-\theta)(b+h-c^t)$, which are both positive. We obtain

$$\frac{\partial \mathbb{E}_p[SW^F(p)]}{\partial \delta} = \frac{c^e}{\delta^2}\Bigg[(\theta C^t - c^e)\frac{f(p_2^F)}{D_2} - ((1-\theta)C^t + c^e)\frac{f(p_1^F)}{D_1} - (\theta C^t + (1-\theta)(b+h-C^t))\frac{\theta c^t - c^e/\delta}{D_2}\frac{f(p_2^F)}{D_2} + ((1-\theta)C^t + \theta(b+h-C^t))\frac{(1-\theta)c^t + c^e/\delta}{D_1}\frac{f(p_1^F)}{D_1}\Bigg]$$

When priors have a uniform distribution, because $f(.) = 1$, the result simplifies into

$$\frac{\partial \mathbb{E}_p[SW^F(p)]}{\partial \delta} = \frac{c^e}{\delta^3 D_1^2 D_2^2}(u + v\delta),$$

where

$$u \triangleq c^e C^t[\theta D_1^2 + (1-\theta)D_2^2] + c^e(b+h-C^t)[(1-\theta)D_1^2 + \theta D_2^2],$$

$$v \triangleq -c^e(b+h)D_1D_2 + (2\theta-1)\theta(1-\theta)(b+h)^2(C^t - c^t)(b+h-2c^t).$$

Because the partial derivative is proportional to a linear expression in δ, it can change sign at most once. Furthermore, the slope is negative if and only if $v < 0$. When the slope is negative, the partial derivative in case (iii) either remains negative, or is positive and then negative as δ increases. If $b+h-2c^t \le 0$, then $v < 0$. When $b+h-2c^t > 0$, we have $v < 0$ if and only if

$$c^e > \frac{(2\theta-1)\theta(1-\theta)(b+h)(C^t - c^t)(b+h-2c^t)}{D_1D_2}.$$

When $b+h-2c^t > 0$ and c^e is less than the above threshold, the slope is positive, so the derivative either remains positive or is first negative then positive as δ increases. *Q.E.D.*

Proof of Proposition 7. Let $A = C^t,\ B = b+h-C^t,\ A' = \delta c^t - \Delta r$ and $B' = \delta(b+h-c^t) + \Delta r$. Then, adjusting parameters $b, h, c^t, \delta, \Delta r$ means we seek A' and B' so that $p_1^F = p_1^S$ and $p_2^F = p_2^S$, that is,

$$\begin{cases} \frac{(1-\theta)A + c^e}{(1-\theta)A + \theta B} = \frac{(1-\theta)A' + c^e}{(1-\theta)A' + \theta B'} \\ \frac{\theta A - c^e}{\theta A + (1-\theta)B} = \frac{\theta A' - c^e}{\theta A' + (1-\theta)B'}. \end{cases}$$

After simplification, we find this system is equivalent to

$$\begin{cases} B - B' = A - A' \\ \theta(1-\theta)(AB' - A'B) = c^e[\theta(B-B') + (1-\theta)(A-A')] \end{cases}$$

$$\Leftrightarrow \quad A - A' = B - B' = 0 \quad \text{or} \quad \begin{cases} B - B' = A - A' \neq 0 \\ c^e = \theta(1-\theta)(b+h-2C^t). \end{cases}$$

$A - A' = B - B' = 0$ means $C^t = \delta c^t - \Delta r$ and $b+h-C^t = \delta(b+h-c^t) + \Delta r$, which is equivalent to $\delta = 1$ and $\Delta r = c^t - C^t$. In this case, it is easy to check that we also have $\bar{c}^F = \bar{c}^S$.

Now consider the other solution, which is possible when $c^e = \theta(1-\theta)(b+h-2C^t)$. In this case, let D such that $D = B - B' = A - A'$. Combining the equations $D = C^t - \delta c^t + \Delta r$ and $D = b + h - C^t - \delta(b+h-c^t) - \Delta r$, it follows that $D = (b+h)(1-\delta)/2$. Plugging into the expression of $\bar{c}^F$, we find that $\bar{c}^F = \bar{c}^S$ implies $\delta = 1$, which implies $D = 0$ and thus reduces to the previous solution. *Q.E.D.*

Proof of Proposition 8. We start with the case in which the physician exerts low effort. In this case, the physician should order a test if and only if

$$pr^+ + (1-p)r^- + \delta pb - \delta c^t > r^n - \delta ph,$$

which is equivalent to

$$p > \frac{\delta c^t - (r^- - r^n)}{\delta(b+h) + r^+ - r^-}.$$

and provide a mild diagnosis otherwise. To ensure $0 \leq \frac{\delta c^t - (r^- - r^n)}{\delta(b+h) + r^+ - r^-} < 1$, we need

$$r^+ - r^n > \delta(c^t - b - h). \tag{A4}$$

If $r^+ - r^n \leq \delta(c^t - b - h)$, the physician never orders a test, regardless of p. Different from the case of the fee-for-service payment scheme, *no* case exists in which the physician orders a test for all patients.

Next, we consider the case in which the physician exerts effort in the consultation process. Depending on the private signal, two scenarios exists:

(i) If the consultation generates a positive signal, the physician updates the patient's probability of suffering from the positive condition:

$$\Pr(s = \bar{s} | \sigma = \bar{\sigma}) = \frac{\theta p}{\theta p + (1-\theta)(1-p)}.$$

The physician should order a test if and only if

$$\frac{\theta p}{\theta p + (1-\theta)(1-p)} \cdot (r^+ + \delta b) + \frac{(1-\theta)(1-p)}{\theta p + (1-\theta)(1-p)} \cdot r^- - \delta c^t > r^n - \delta \cdot \frac{\theta p}{\theta p + (1-\theta)(1-p)} \cdot h,$$

which gives

$$p > \frac{(1-\theta)[\delta c^t - (r^- - r^n)]}{(1-\theta)[\delta c^t - (r^- - r^n)] + \theta[\delta(b+h-c^t) + r^+ - r^n]}.$$

To ensure $0 \leq \frac{(1-\theta)[\delta c^t - (r^- - r^n)]}{(1-\theta)[\delta c^t - (r^- - r^n)] + \theta[\delta(b+h-c^t) + r^+ - r^n]} \leq 1$, we need $r^+ - r^n > \delta(c^t - b - h)$, which coincides with (A4). If $r^+ - r^n \leq \delta(c^t - b - h)$, the physician never orders a test, regardless of p. No case exists in which the physician orders a test for all patients.

(ii) If the consultation generates a positive signal, the physician updates the patient's probability of suffering from the positive condition:

$$\Pr(s = \bar{s} | \sigma = \underline{\sigma}) = \frac{(1-\theta)p}{(1-\theta)p + \theta(1-p)}.$$

The physician should order a test if and only if

$$\frac{(1-\theta)p}{(1-\theta)p + \theta(1-p)} \cdot (r^+ + \delta b) + \frac{\theta(1-p)}{(1-\theta)p + \theta(1-p)} \cdot r^- - \delta c^t > r^n - \delta \cdot \frac{(1-\theta)p}{(1-\theta)p + \theta(1-p)} \cdot h,$$

which is equivalent to

$$p > \frac{\theta[\delta c^t - (r^- - r^n)]}{\theta[\delta c^t - (r^- - r^n)] + (1-\theta)[\delta(b+h-c^t) + r^+ - r^n]}.$$

To ensure $0 \leq \frac{\theta[\delta c^t - (r^- - r^n)]}{\theta[\delta c^t - (r^- - r^n)] + (1-\theta)[\delta(b+h-c^t) + r^+ - r^n]} < 1$, we need $r^+ - r^n > \delta(c^t - b - h)$, which coincides with (A4). If $r^+ - r^n \leq \delta(c^t - b - h)$, the physician never orders a test, regardless of p. No case exists in which the physician orders a test for all patients.

Finally, we compare the two cases (i.e., the physician exerts low vs. high effort) and analyze the physician's effort decision.

(i) If $r^+ - r^n \leq \delta(c^t - b - h)$, the physician should choose a low effort level.

(ii) If $r^+ - r^n > \delta(c^t - b - h)$, the physician's effort decision can be determined as follows:

(1) If $\frac{\theta[\delta c^t - (r^- - r^n)]}{\theta[\delta c^t - (r^- - r^n)] + (1-\theta)[\delta(b+h-c^t) + r^+ - r^n]} < p \leq 1$, the expert always orders the test regardless of the effort or the signal. Thus, the expert exerts low effort.

(2) If $0 < p \leq \frac{(1-\theta)[\delta c^t - (r^- - r^n)]}{(1-\theta)[\delta c^t - (r^- - r^n)] + \theta[\delta(b+h-c^t) + r^+ - r^n]}$, the expert always provides a mild diagnosis regardless of the effort or the signal. Thus, the expert exerts low effort.

(3) If $\frac{(1-\theta)[\delta c^t - (r^- - r^n)]}{(1-\theta)[\delta c^t - (r^- - r^n)] + \theta[\delta(b+h-c^t) + r^+ - r^n]} < p \leq \frac{\theta[\delta c^t - (r^- - r^n)]}{\theta[\delta c^t - (r^- - r^n)] + (1-\theta)[\delta(b+h-c^t) + r^+ - r^n]}$, the expert follows the signal if exerting effort. The condition for the physician to exert effort is as follows:

(a) If $\frac{(1-\theta)[\delta c^t - (r^- - r^n)]}{(1-\theta)[\delta c^t - (r^- - r^n)] + \theta[\delta(b+h-c^t) + r^+ - r^n]} < p \leq \frac{\delta c^t - (r^- - r^n)}{\delta(b+h) + r^+ - r^-}$, the physician exerts effort if and only if

$$\begin{aligned}
&[\theta p + (1-\theta)(1-p)] \cdot \left[\frac{\theta p}{\theta p + (1-\theta)(1-p)} \cdot (r^+ + \delta b) + \frac{(1-\theta)(1-p)}{\theta p + (1-\theta)(1-p)} \cdot r^- - \delta c^t\right] \\
&+ [(1-\theta)p + \theta(1-p)] \cdot \left[r^n - \delta \cdot \frac{(1-\theta)p}{(1-\theta)p + \theta(1-p)} \cdot h\right] - c^e \\
&> r^n - \delta p h,
\end{aligned}$$

which is equivalent to

$$p > \frac{(1-\theta)[\delta c^t - (r^- - r^n)] + c^e}{(1-\theta)[\delta c^t - (r^- - r^n)] + \theta[\delta(b+h-c^t) + r^+ - r^n]}.$$

We can easily verify that

$$\frac{(1-\theta)[\delta c^t - (r^- - r^n)] + c^e}{(1-\theta)[\delta c^t - (r^- - r^n)] + \theta[\delta(b+h-c^t) + r^+ - r^n]} > \frac{(1-\theta)[\delta c^t - (r^- - r^n)]}{(1-\theta)[\delta c^t - (r^- - r^n)] + \theta[\delta(b+h-c^t) + r^+ - r^n]}.$$

To ensure

$$\frac{(1-\theta)[\delta c^t - (r^- - r^n)] + c^e}{(1-\theta)[\delta c^t - (r^- - r^n)] + \theta[\delta(b+h-c^t) + r^+ - r^n]} \leq \frac{\delta c^t - (r^- - r^n)}{\delta(b+h) + r^+ - r^-},$$

we need

$$c^e \leq \bar{c}^d \triangleq \frac{(2\theta - 1)[\delta c_t - (r^- - r^n)][\delta(b+h-c_t) + r^+ - r^n]}{\delta(b+h) + r^+ - r^-}. \tag{A5}$$

If c^e is above the threshold, the physician exerts low effort for any patient within this range of priors.

(b) If $\frac{\delta c^t-(r^- - r^n)}{\delta(b+h)+r^+-r^-} < p \le \frac{\theta[\delta c^t-(r^- - r^n)]}{\theta[\delta c^t-(r^- - r^n)]+(1-\theta)[\delta(b+h-c^t)+r^+-r^n]}$, the physician exerts effort if and only if

$$\begin{aligned}&[\theta p+(1-\theta)(1-p)]\cdot\left[\frac{\theta p}{\theta p+(1-\theta)(1-p)}\cdot(r^+ + \delta b)+\frac{\theta p}{\theta p+(1-\theta)(1-p)}r^- - \delta c^t\right]\\&+[(1-\theta)p+\theta(1-p)]\cdot\left[r^n-\delta\cdot\frac{(1-\theta)p}{(1-\theta)p+\theta(1-p)]}\cdot h\right]-c^e\\&> pr^+ + (1-p)r^- + \delta pb - \delta c^t,\end{aligned}$$

which yields

$$p<\frac{\theta[\delta c^t-(r^- - r^n)]-c^e}{\theta[\delta c^t-(r^- - r^n)]+(1-\theta)[\delta(b+h-c^t)+r^+-r^n]}.$$

We can easily verify that

$$\frac{\theta[\delta c^t-(r^- - r^n)]-c^e}{\theta[\delta c^t-(r^- - r^n)]+(1-\theta)[\delta(b+h-c^t)+r^+-r^n]}<\frac{\theta[\delta c^t-(r^- - r^n)]}{\theta[\delta c^t-(r^- - r^n)]+(1-\theta)[\delta(b+h-c^t)+r^+-r^n]}.$$

To ensure

$$\frac{\theta[\delta c^t-(r^- - r^n)]-c^e}{\theta[\delta c^t-(r^- - r^n)]+(1-\theta)[\delta(b+h-c^t)+r^+-r^n]}\ge\frac{\delta c^t-(r^- - r^n)}{\delta(b+h)+r^+-r^-},$$

we need

$$c^e\le\frac{(2\theta-1)[\delta c_t-(r^- - r^n)][\delta(b+h-c_t)+r^+-r^n]}{\delta(b+h)+r^+-r^-}=\bar{c}^d,$$

which is identical to (A5). If c^e is above the threshold, the physician exerts low effort for any patient within this range of priors. *Q.E.D.*

Proof of Proposition 9. Setting $p_1^d = p_1^S$ and $p_2^d = p_2^S$ gives

$$\begin{aligned}&r^+ - r^n = b - C^t + h - \delta(b-c^t+h) = \delta c^t - C^t + (1-\delta)(b+h)\\&r^- - r^n = \delta c^t - C^t.\end{aligned}$$

We can verify that under the above solution,

$$\begin{aligned}\bar{c}^d&=\frac{(2\theta-1)[\delta c_t-(r^- - r^n)][\delta(b+h-c_t)+r^+-r^n]}{\delta(b+h)+r^+-r^-}\\&=\frac{C^t(2\theta-1)(b+h-C^t)}{b+h},\end{aligned}$$

which is equal to $\bar{c}^S$. The proof is complete. *Q.E.D.*

A3: Alternate Definition of Social Welfare

In this section, we consider a definition of social welfare that differs from that used in the main body. Specifically, we define an alternative social welfare as the combination of expected patient utility, payer expenditure, and physician utility:

$$SW = \Pi_{\text{payer}} + U_{\text{physician}} + \mathbb{E}[U_{\text{patient}}].$$

This definition contrasts with that used in the main body of the paper, in that we use as the second term of social welfare the physician utility $U_{\text{physician}} = \Pi_{\text{physician}} + \delta\mathbb{E}[U_{\text{patient}}]$ instead of physician's payoff $\Pi_{\text{physician}}$. This new definition of social welfare implies

$$SW = \Pi_{\text{payer}} + \Pi_{\text{physician}} + (\delta+1)\mathbb{E}[U_{\text{patient}}];$$

that is, the expected patient utility has a weight $\delta+1$, whereas in the main body of the paper, this term has a coefficient equal to 1, which changes the weight of parameters b, h, and c^t in social welfare. More fundamentally, the weight of the patient utility in social welfare $(\delta+1)$ is now guaranteed to exceed the weight of the patient utility in the physician's objective (δ). By contrast, in our base model, the weight of the patient utility in social welfare is equal to 1, which could be lower than the weight of the patient utility in the physician's objective (if $\delta > 1$).

The results focused on the physician's optimal policy are unaffected. Only the results that make use of the socially optimal policy are affected. Namely, Proposition 2 is modified as follows (the proofs of all modified results are omitted for the sake of brevity; they are very similar to the proof of the corresponding results presented in Appendix A2):

Let

$$\bar{c}^S \triangleq (2\theta-1)(C^t+\delta c^t)\left(1-\frac{C^t+\delta c^t}{(\delta+1)(b+h)}\right)$$

$$p_1^S \triangleq \frac{(1-\theta)(C^t+\delta c^t)+c^e}{(1-\theta)(C^t+\delta c^t)+\theta((\delta+1)(b+h)-C^t-\delta c^t)}$$

$$p_2^S \triangleq \frac{\theta(C^t+\delta c^t)-c^e}{\theta(C^t+\delta c^t)+(1-\theta)((\delta+1)(b+h)-C^t-\delta c^t)}.$$

Proposition A5. *The socially optimal policy is as follows:*

(i) If $(\delta+1)(b+h) \le C^t+\delta c^t$, the physician exerts low effort and does not order a test for any patient.

(ii) If $(\delta+1)(b+h) > C^t+\delta c^t$ and $c^e > \bar{c}^S$, the physician exerts low effort for all patients and orders a test if and only if $p \ge (C^t+\delta c^t)/((\delta+1)(b+h))$.

(iii) If $(\delta+1)(b+h) > C^t+\delta c^t$ and $c^e \le \bar{c}^S$, the socially optimal policy depends on the patient's prior:

(a) if $p \le p_1^S$, the physician exerts low effort and does not order a test;

(b) if $p_1^S < p \le p_2^S$, the physician exerts high effort and follows the signal (i.e., if the signal is positive, order a test; if the signal is negative, do not order a test);

(c) if $p > p_2^S$, the physician exerts low effort and orders a test.

Although the thresholds for the different cases are slightly modified, the general structure of the socially optimal policy remains unchanged.

Proposition 4 is modified as follows:

Proposition A6. *When Δr is such that the range of priors with high effort is at its widest under the fee-for-service physician's optimal strategy, that range is wider than the socially optimal range if and only if either $C^t+\delta c^t < c_0$ or $C^t+\delta c^t > c_1$, where*

$$c_0 \triangleq \frac{(\delta+1)(b+h)}{2}(1-\sqrt{q}); \quad c_1 \triangleq \frac{(\delta+1)(b+h)}{2}(1+\sqrt{q}), \quad q \triangleq \frac{c^e}{(2\theta-1)(1+\delta)[c^e(2\theta-1)+\theta(1-\theta)\delta(b+h)]}.$$

Otherwise (i.e., $c_0 \le C^t+\delta c^t \le c_1$), the socially optimal range of high effort is wider than the fee-for-service range of high effort for all Δr.

The definitions of c_0 and c_1 are modified, but the main difference from the result in the base model is that $\delta > 1$ no longer guarantees fee-for-service leads to a wider range of priors leading to high effort than the social optimum. The reason is that in the base model, $\delta > 1$ corresponds to a situation where the social planner weighs the patient utility less than the physician—a case that can no longer occur with this alternate definition. However, the insight that fee-for-service may lead to a wider range of priors leading to high effort than the social optimum remains true. It would be the case for sufficiently low or sufficiently high cost of the test (condition similar to that obtained in the base model for $\delta \le 1$).

The intermediate result Proposition A3 is modified as follows:

Proposition A7. *Under fee-for-service, the average population social welfare is as follows:*
(i) If $\Delta r \le \delta(c^t - b - h)$*, then* $\mathbb{E}_p[SW^F(p)] = -(\delta+1)h\mu$*;*
(ii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ *and* $c^e > \bar{c}^F$*, then*

$$\mathbb{E}_p[SW^F(p)] = b(\delta+1)\mu - (C^t + \delta c^t)\bar{F}\left(\frac{\delta c^t - \Delta r}{\delta(b+h)}\right) - (b+h)(\delta+1)Q\left(\frac{\delta c^t - \Delta r}{\delta(b+h)}\right);$$

(iii) If $\delta(c^t - b - h) < \Delta r < \delta c^t$ *and* $c^e \le \bar{c}^F$*, then*

$$\begin{aligned}\mathbb{E}_p[SW^F(p)] = {} & b(\delta+1)\mu - C^t - \delta c^t - c^e(F(p_2^F) - F(p_1^F)) + (C^t + \delta c^t)(\theta F(p_2^F) + (1-\theta)F(p_1^F)) \\ & - (C^t + \delta c^t)\left(\theta Q(p_2^F) + (1-\theta)Q(p_1^F)\right) - ((b+h)(\delta+1) - C^t - \delta c^t)\left((1-\theta)Q(p_2^F) + \theta Q(p_1^F)\right);\end{aligned}$$

(iv) If $\Delta r \ge \delta c^t$*, then* $\mathbb{E}_p[SW^F(p)] = (\delta+1)b\mu - C^t - \delta c^t$*.*

Using this intermediate result, Proposition 5 is modified as follows:

Proposition A8. *If* $c^e > (2\theta - 1)\delta(b+h)/4$*, the value of* Δr *that maximizes the average population social welfare under fee-for-service is*

$$\Delta r = \begin{cases} -\delta(C^t - c^t)/(\delta+1) & \text{if } (\delta+1)(b+h) > C^t + \delta c^t \\ \text{any value within } (-\infty, -\delta(b+h-c^t)] & \text{otherwise.} \end{cases}$$

The above result is similar to that in the base model (with slightly modified optimal Δr and associated condition) and leads to the same insights.

Lemma 6 is modified as follows: suppose $\Delta r = 0$ and the patient priors are uniformly distributed. For δ below a threshold, social welfare is monotonic with respect to δ, with a slope equal to $b/2 - c^t + (c^t)^2/(2(b+h))$, which is positive if $b + h - c^t > \sqrt{h(b+h)}$, and negative otherwise. This property means social welfare may be decreasing in δ in this region (when $0 < b + h - c^t < \sqrt{h(b+h)}$). Above the threshold, analytically studying the sign of the derivative of social welfare with respect to δ is intractable: the derivative can no longer be written as a linear expression of δ divided by δ^3. Instead, it is a polynomial of degree 3 divided by δ^3. However, we find numerically that in this region, social welfare may also be decreasing with δ.

Let us now focus on the incentive-alignment results (in Section 6.4). Proposition 9 is modified as follows:

Proposition A9. *The physician's effort and testing decisions maximize social welfare under a diagnosis-based payment scheme with* $r^- - r^n = -C^t$ *and* $r^+ - r^n = b + h - C^t$*. In particular,* $r^- - r^n < 0$ *and, if* $b + h > C^t$*, then* $r^+ - r^n > 0$*.*

The result indicates that, even with the alternate definition of social welfare, a diagnosis-based payment scheme can align the physician's decisions to the social optimum. The main difference from the result obtained in the main body of the paper is that there is no longer any condition on δ to ensure $r^+ - r^n > 0$.