Pricing Diagnosis-Based Services When Customers Exhibit Sunk Cost Bias

Guangwen Kong
Industrial & Systems Engineering Department, University of Minnesota, Minneapolis, MN 55455, USA
Email: gkong@umn.edu; Phone: (612) 624-6610

Sampath Rajagopalan
Marshall School of Business, University of Southern California, Los Angeles, CA 90089, USA
Email: raj@marshall.usc.edu; Phone: (213) 740-0193

Chunyang Tong
School of Economics and Management, Tongji University, Shanghai, 200092, China
Email: tongchunyang@tongji.edu.cn; Phone: (86)21-6598-1473

Abstract

Significant evidence has emerged that consumers are boundedly rational and display a sunk cost bias when making decisions. Moreover, customers may display dynamically inconsistent beliefs about the extent to which they will discount sunk costs. We evaluate the impact of sunk cost bias and dynamically inconsistent preferences on commonly used pricing schemes in a service setting with diagnosis and treatment phases. Using a two-part tariff pricing scheme, a monopolist service provider can achieve higher profits by exploiting the dynamic inconsistency of customers in their sunk cost bias proclivity. In fact, the provider can achieve higher profits when sunk cost bias is higher or the dynamic inconsistency is greater. In contrast, the provider’s profit is lower when sunk cost bias is higher under a single rate time-based fee. We evaluate special cases of the two-part tariff frequently used in practice such as conditional diagnosis fee, single rate time-based fee, and free diagnosis and characterize their relative performance under different levels of bounded rationality. A conditional diagnosis fee, which is charged only if a customer leaves after diagnosis without treatment, does very well. Offering a free diagnosis and charging a high rate for treatment performs poorly. We also provide insights into a setting in which the market consists of two customer classes with different levels of sunk cost bias or
dynamic inconsistency in preferences. Overall, we show that sunk cost bias and inconsistency in beliefs over time are important factors that a provider should carefully consider in choosing a pricing scheme.

**Keywords:** Pricing of Services, Bounded rationality, Sunk cost bias, Diagnostic services

**Received:** December 2016; **Accepted:** February 2018 by Haresh Gurnani after three revisions.

1 Introduction

Many services have a diagnostic phase followed by a “treatment” phase. Some examples are equipment diagnosis/repair and consulting. Consulting firms typically perform upfront project scoping prior to starting work on an in-depth consulting project (Salmon and Rosenblatt 1995). The initial phase of the project involves a preliminary problem diagnosis based on discussion with the client and identifying the project scope, which may take a significant amount of time and cost thousands of dollars. The second phase of the project involves detailed fact finding, identification of solutions and recommendations – henceforth called “treatment” phase. When customers are charged an initial diagnosis fee (depending on the pricing scheme chosen) followed by a separate treatment fee, they are likely to mentally include the diagnosis fee in deciding whether to go forward with the treatment, even though the diagnosis fee is a sunk cost and should not influence their decision to continue with treatment. This effectively lowers the amount they are willing to pay for treatment. There is considerable evidence for the prevalence of the sunk cost bias and bounded rationality among consumers (Thaler 1985).

In this work, we explore the impact of bounded rationality on various commonly used pricing schemes by a monopolist provider of a diagnosis-based service. We first consider a two-part tariff comprised of a fixed diagnostic fee and a time-based fee for treatment. In addition, we consider variants and special cases of the two-part tariff, such as a simple fixed fee and time-based pricing based on a single hourly rate for both diagnosis and treatment. Based on an extensive survey, Lowendahl (2005) suggests that the two most commonly used pricing schemes in real-world service settings such as consulting, legal, and repair services are fixed fee and time-based fee schemes. A novel feature of our work is that customers are boundedly rational along two dimensions that impacts their decisions about continuing with the service after diagnosis under some pricing schemes. First, such customers take into account sunk costs incurred when they decide whether to continue with treatment. Second, they may display dynamically inconsistent beliefs about their sunk cost bias before and after diagnosis. That is, a customer may believe that she does not have sunk cost bias before diagnosis, but after paying the diagnosis fee she may actually deduct a fraction of the sunk costs in deciding whether to move forward with treatment. We model this second dimension by
allowing customers to exhibit varying levels of “sophistication” in their prior beliefs about whether they exhibit sunk cost bias. At one extreme, “sophisticated” customers are fully aware (before diagnosis) that they will discount sunk costs after diagnosis when deciding whether to continue with the service. Naive customers, on the other hand, are completely unaware (before diagnosis) that they will discount sunk costs, but in fact do discount sunk costs after diagnosis. This distinction between naive and sophisticated customers is based on the discussion in Spiegler (2011), which provides a comprehensive review of how firms exploit bounded rationality. Sophisticated customers are irrational but consistent in their preferences over time, while naive customers are irrational and inconsistent. Koszegi (2014), DellaVigna and Malmendier (2004), and Wang and Yang (2010) all provide substantial justification for the phenomenon of dynamically inconsistent beliefs. We consider a more general setting with partially sophisticated customers whose beliefs regarding the fraction of the sunk cost they will discount is different before and after diagnosis. We investigate how this subtle difference in beliefs among customers with varying levels of sophistication impacts the attractiveness of different pricing schemes. We first consider a setting in which all customers are homogeneous in these two dimensions of bounded rationality and later a setting in which there is heterogeneity in the customers’ sunk cost bias and sophistication level.

The service process in our model has two phases: a diagnosis phase followed by a treatment phase. Customers are ex-ante homogeneous in terms of future treatment time before the diagnosis phase. In the diagnosis phase, the service provider identifies the service type of each customer – i.e., the time needed to solve the customer’s problem. In the treatment phase, whose duration depends on the service type, the provider solves the customer’s problem and completes the service, which provides a fixed value to the customer. For example, an auto repair shop will first diagnose a problem and then fix the problem during the treatment phase. The value from the service is only realized when the service is completed by the provider, for example, after a consultant report is generated or equipment is repaired. The provider incurs a service cost that varies linearly with time. Services with a diagnostic and treatment phase are attractive settings in which to study the sunk cost bias because a customer has to decide whether to obtain treatment after diagnosis and may exhibit a sunk cost bias in making this decision. The focus of our work is on analyzing the service provider’s pricing strategy and comparing some commonly used pricing schemes when there are boundedly rational consumers, as well as exploring how sunk cost bias and dynamic inconsistency in preferences impacts the pricing schemes.

The main insights from the study are as follows. First, the two-part tariff scheme is effective in exploiting dynamic inconsistency in customer beliefs due to its flexibility in charging different rates for diagnosis and treatment. Hence, such a scheme can overcome the lower willingness to pay engendered by the sunk cost bias and achieve higher profits when consumers are more naive.
Moreover, profits are higher when naive consumers have greater sunk cost bias. This is not the case when customers show no naivete and the provider makes the same profit as in the setting in which all customers are rational, that is, exhibit no sunk cost bias.

We also consider various special cases of the two-part tariff. The provider’s profit under a single rate time-based scheme is adversely impacted by the sunk cost bias unlike under a two-part tariff. However, he can extract more profits from naive customers than from sophisticated ones, as was true under the two-part tariff and for the same reason, viz., the provider exploits the dynamic inconsistency of naive customers. The fixed fee outperforms (underperforms) the time-based fee when the value derived from the service is significantly higher (lower) than the cost of providing the service. When this ratio of value to cost is neither too high nor too low, there exists a threshold on the level of sunk cost bias such that time-based pricing is preferred at lower levels of sunk cost bias and a fixed fee is preferred otherwise. The conditional fixed fee, wherein the provider charges a fixed fee to all treated customers but can turn away a fraction of the customers after diagnosis (those with high treatment time), performs very well and can achieve the same profit as the two-part tariff when customers are sophisticated. It also outperforms the simple fixed fee and time-based fee. Many repair shops charge a diagnosis fee only if the customer leaves after diagnosis without treatment. Such a conditional two-part tariff achieves the same profit as a conditional fixed fee. Another scheme seen in practice, also a special case of the two-part tariff, is to offer a free diagnosis to attract customers and possibly charge a high rate for treatment. We show that such a scheme is largely dominated by the other pricing schemes and can perform poorly.

When the population is heterogeneous and comprised of two types with different levels of sunk cost bias, we find that it is not more profitable for the provider to offer two different two-part tariff contracts to serve both segments. Instead, it is optimal to either serve both types with a contract that is optimal in serving only the type with a higher sunk cost bias or only the type with a lower sunk cost bias. A similar conclusion holds in the case in which customers are heterogeneous in their sophistication level rather than in their sunk cost bias. When customers are heterogeneous in both dimensions of bounded rationality, the optimal two-part tariff contract offered and the segments that are served depend on various problem parameters. We also find that the provider should not ignore the heterogeneity in the customer base unless the proportion of customers with high sunk cost bias and sophistication level is low.

1.1 Literature review

The sunk cost bias is a form of bounded rationality and has been extensively studied in the realm of psychology and behavioral economics. Thaler (1985) proposes and coins the notion of “mental accounting”, that is, consumers make their joining decisions with a mental account of their expenses.
that influences their future consumption. Tversky and Kahneman (1981) use the theory of framing to explain this psychological bias. Heath (1995) shows that consumers may be more or less likely to continue with a task after having expended some initial effort or cost, depending on whether the initial investments are easy to keep track of and mental budgets are easy to create. Prelec and Loewenstein (1998) investigate how mental accounting influences behaviors with respect to saving and debt. More specifically, Arkes and Ayton (1999) explain the sunk cost bias under which people tend to incorporate their previous payment when valuing prospective consumption. Dick and Lord (1998) examine how a consumer’s behavior changes if there exists a sunk cost bias related to the upfront payment – for instance, a membership fee. Gourville and Soman (1998) identify customers’ bias of “payment depreciation” when the payment is temporally separated from consumption. Horngren et al. (2000) reports on surveys in which a majority of managers in the U.S., U.K, and Australia take sunk costs into account when making pricing decisions. Recently, Herrmann et al. (2015) empirically show a strong sunk cost bias in an auction setting and Ho et al. (2017) provide strong evidence of sunk cost bias among car buyers in Singapore and Hong Kong.

Another stream of literature explores how firms may take advantage of bounded rationality in making pricing and product differentiation decisions. Narasimhan et al. (2005) is an early review paper urging marketing researchers to incorporate behavioral anomalies in their theoretical models to determine implications for a firm’s strategic decisions. Ho et al. (2006) explain in detail how ideas from behavioral economics can be incorporated into various marketing models, including those focused on pricing decisions. Spiegler (2011) provides a comprehensive review of how firms may exploit customers’ bounded rationality, including dynamically inconsistent preferences, reference dependency effects, and so forth. Goldfarb et al. (2012) provide a recent review of the literature that focuses on the impact of biases and bounded rationality on managers, rather than on consumers. Soman (2001) examines how customer perceptions of transaction decoupling impacts a firm’s price bundling decision. Popescu and Wu (2007) investigate firms’ dynamic pricing strategy when customers exhibit reference effects. Lim and Ho (2007) study the price contract design when customers are boundedly rational in a manufacturer-retailer channel setting. Cui et al. (2007) consider fairness concerns in supply chain contracts and show that a simple wholesale price can achieve supply chain coordination. Ho and Zhang (2008) study channel efficiency in the presence of retailers’ bounded rationality in the form of reference dependence. Similar to our study, they consider a scenario wherein an up-front fixed payment from players can be perceived as a loss. Erat and Bhaskaran (2012) conduct experiments to show that consumers’ willingness to pay for complementary add-ons is correlated with the price of the base product. They further investigate how this consumer bias influences a firm’s pricing strategies but do not endogenize the choice of a pricing scheme in a service setting, as we do in the current paper. Jiang et al. (2017) study the effects of
anticipated regret on product innovation and a firm’s profit in both monopolistic and competitive settings. However, these works do not study the choice of pricing scheme in the presence of a sunk cost bias in a B2C service setting. Lambrecht et al. (2011) provide a recent review of pricing and price discrimination in service industries. They point out that one avenue for future research is the study of dynamic environments where consumers do not always make rational choices due to their time-inconsistency and our work considers such a setting.

This work is also related to the literature on fixed and time-based pricing of services. Anand et al. (2011) consider pricing of discretionary services in which the value or quality derived by a customer depends upon the service time in an endogenous demand model, but only using a fixed fee. Tong and Rajagopalan (2014) analyze the optimal choice of fixed and time-based pricing of discretionary services when the variation in service time is endogenous. Zhang et al. (2017) analyze the optimal choice of a fixed fee and a time-based fee in professional services in which customers have heterogeneous valuations and the provider’s effort is not observable. Please refer to Tong and Rajagopalan (2014) and Zhang et al. (2017) for additional references to this literature. Prior literature in this area typically assumes that customers are fully rational and does not consider sunk cost bias.

There is a growing literature in operations management on the effects of bounded rationality. This literature has explored the impact of bounded rationality on ordering and inventory decisions by a single decision maker as well as on supply chain contracts in distribution channels. Gino and Pisano (2008) argue persuasively for the need to incorporate departures from rationality in the models developed in the operations management literature, as has been done in other areas such as marketing and economics. Our work is among the few papers investigating customers’ bounded rationality or behavioral bias in a service setting. Huang et al. (2013) explore bounded rationality, manifested as a customer’s inability to accurately estimate waiting times, in a service setting. They explore the impact on a firm’s pricing decision in both visible and invisible queues and highlight the importance of incorporating bounded rationality in pricing decisions. Huang and Chen (2015) consider a firm’s pricing strategy when customers can only rely on past experiences and anecdotal reasoning to make their joining decisions. Plambeck and Wang (2013) explore consumers’ bounded rationality in undergoing a service experience that is unpleasant in the short run but has long-term benefits. They consider a firm’s pricing and scheduling decisions in such a system and analyze the attractiveness of usage versus subscription fees in various scenarios. The above-mentioned works mainly focus on a simple pricing scheme. One exception is Katok et al. (2016), which considers coordination behavior between contractors under risk-sharing contracts in project management. We consider different types of pricing schemes when facing customers with sunk cost bias.

While there is an extensive and growing literature on pricing and bounded rationality as previ-
ously described, the impact of a sunk cost bias on a firm’s pricing decisions has not been studied, as indicated in Bendoly et al. (2006). This is the case despite the importance of this issue in many service settings where there is a diagnostic phase, for which a fee is paid, followed by a “treatment” phase. To the best of our knowledge, we are the first to study how the sunk cost bias impacts the choice and nature of the pricing scheme offered by a monopolist.

2 Model Primitives

A provider sells a service that requires a diagnosis phase followed by a “treatment” or service delivery phase. As discussed previously, many services such as consulting, legal, and equipment repair have such a two-phase structure. The customer derives a value $v_0$ only upon satisfactory completion of the service after the service delivery phase, which is the value that follows from resolution of the problem. For example, a customer derives value from the use of repaired equipment and this value is independent of the nature of the repair and associated cost. If a customer aborts the service after diagnosis, the value derived is assumed to be 0 because the value is derived only if the problem is resolved. We assume that the value $v_0$ derived from the service is homogeneous and is independent of the treatment time.$^1$

The time required for diagnosis is fixed and equal to $\tau_0$; the service delivery time $\tau$ is uncertain and depends upon the diagnosis.$^2$ For example, a consultant will diagnose the issues facing a client, which will provide an estimate of the time it will take to provide a detailed analysis and recommendation. Thus, the provider needs to spend time $\tau_0$ to diagnose the problem, which in turn will determine how much additional time $\tau$ will be required to provide “treatment” to solve the problem. A similar approach to modeling a diagnostic service has been adopted by Jiang et al. (2014). We refer to the service time $\tau \in [0,1]$ as the customer type, which becomes known only after diagnosis to both the customer and the provider. In other words, to achieve a service outcome with value $v_0$, the provider will need to spend a total of $(\tau_0 + \tau)$ amount of time, with $\tau_0$ being the time required for discovering a customer’s type.$^3$ We assume that the customer type $\tau$ is uniformly distributed with support $[0,1]$. Other works in the literature have made similar assumptions, see Lane (1980), Mendelson and Parlakturk (2008), and more recently Fuchs et al. (2015). The provider

$^1$However, the results and insights are applicable to a setting where customer’s value increases with treatment time as explained in footnote 4.

$^2$Although it is easy to extend the model to a scenario with random diagnosis time $\tau_0$, this increases the notational complexity without adding any new insights.

$^3$To isolate and focus exclusively on the effect of irrational customer behavior (specifically the sunk cost bias) on the choice of pricing scheme by the provider, we assume that the provider does not misreport the problem and service time. The impact of asymmetric information on professional services has been well documented in the literature (interested readers are referred to the comprehensive review by Dulleck and Kerschbamer (2006)). In addition, customers can prevent such misreporting using advanced technology to monitor such activities or by consulting another service provider. Also, such misreporting is less likely given the availability of online reviews and reputation concerns, when consumers can easily influence the reputation of even small service providers using services such as Yelp.
incurs a linear cost \( c \) per unit time to provide the service. This cost may represent hourly wages paid to employees or simply the opportunity cost of the provider’s time. The provider may also incur some material costs as part of the treatment. Such costs can be easily incorporated but do not impact the choice of pricing schemes in our study and are therefore ignored.

A provider can potentially choose from a variety of pricing schemes. We first consider a two-part tariff scheme comprised of a fixed fee for diagnosis and a time-based fee for treatment. Suppose the provider charges a fixed fee \( f \) for diagnosis and a rate \( r \) per unit time for treatment if she decides to continue with the service. Thus, a customer will pay \( f \) when starting the service and pay \( r \tau \) after diagnosis based on the customer type \( \tau \) identified. Such a two-part tariff is also equivalent to charging two different rates for diagnosis and treatment because the diagnosis time is fixed. The two-part tariff is a flexible scheme and other commonly used schemes, such as fixed fee or time-based pricing (with the same rate for both diagnosis and treatment), are special cases of the two-part tariff. Another special case of the two-part tariff is when the diagnosis is free and a time-based fee for treatment is charged. Wong (2014) provides some persuasive reasons why simple pricing schemes are used. We initially focus on the two-part tariff pricing scheme and later discuss special cases of the two-part tariff.

2.1 Bounded Rationality

In our setting, customers are boundedly rational along two dimensions. First, unlike a rational customer, customers take into account sunk costs incurred during the diagnosis stage when they decide to continue with treatment. Second, they may display dynamically inconsistent preferences in their sunk cost bias before and after diagnosis. Both of these aspects impact a consumer’s evaluation of a pricing scheme and decision to continue with the service after diagnosis. We describe these two modeling features in detail next. First, we consider how customers account for the sunk cost \( f \) at time \( \tau_0 \) after diagnosis. Specifically, at time \( \tau_0 \), if identified as type \( \tau \), a customer will continue with the service if and only if \( r \tau \leq v_0 - f \theta \), where \( \theta \in [-1, 1] \) measures the degree of sunk cost bias. Notice that \( \theta = 0 \) represents rational customers who do not have sunk cost bias – i.e., they do not factor in sunk costs when deciding on treatment. Henceforth, we use the phrase “discount sunk cost” to refer to this phenomenon. Unlike rational customers, the boundedly rational customers include a fraction \( \theta \) of the sunk cost \( f \) in comparing the costs and benefits of the service at time \( \tau_0 \). The case \( \theta > (\leq) 0 \) represents a scenario in which a customer, after diagnosis, is willing to pay less (more) for future consumption in the presence of a sunk cost bias. We initially assume that customers are homogeneous in their sunk cost bias \( \theta \) and later extend the model to a market with two customer types with different \( \theta \) values.

While some prior studies have suggested that incurring a sunk cost (diagnosis cost in our case)
may escalate commitment to a service (i.e., $\theta < 0$), Heath (1995) found in his experimental study that escalating commitment is not a universal reaction to sunk costs. When individuals were given explicit information on total income and total costs, errors of de-escalation occurred. Subjects used “mental accounting” and were reluctant to continue investing when total costs exceeded total benefits, even when such a choice was not optimal. In fact, Heath (1995) (page 39), concluded that when “initial investment is easy to keep track of and mental budget is easy to create,” a customer’s willingness to pay for future consumption decreases in the presence of sunk costs, i.e., $\theta > 0$. This is also consistent with the mental accounts approach, wherein a customer has a negative mental account based on the diagnosis fee paid (Thaler (1985)) and has to incur additional costs before deriving the value. More recently, Soman (2001) provides further support for the findings of Heath (1995) by showing that past payments strongly reduce purchase intention when the consumer’s wealth is depleted immediately rather than later. Since the nature of diagnosis-based service in our study has this characteristic, we focus on the case $\theta > 0$.

The second dimension of bounded rationality we capture is the extent to which customers display dynamically inconsistent preferences. In Spiegler (2011), a book devoted to pricing models when consumers are boundedly rational, classifies consumers into two types: sophisticated and naive. Although both sophisticated and naive customers suffer from bounded rationality, the difference between them is that sophisticated customers are aware of their bounded rationality while naive customers are not. Similar distinctions between sophisticated and naive customers are used in a number of recent articles such as Armstrong and Vickers (2012), Gu and Wenzel (2014) and Kosfeld and Schuwer (2011). We consider a more general setting wherein consumers may exhibit partial sophistication and capture it through a parameter $s \in [0, 1]$, which represents a customer’s sophistication level. In our setting, this implies that the customer assumes at time 0 that she will discount the sunk cost by an amount $fs\theta$ at time $\tau_0$ in making a decision. However, after the diagnosis, when she is about to make a decision at time $\tau_0$ about continuing with the service, the customer will in fact discount the sunk cost by $f\theta$. Thus, the sophisticated customer ($s = 1$) knows at time 0 that she will discount the sunk cost by $f\theta$ at time $\tau_0$ and is consistent at time $\tau_0$ in her discounting. On the other hand, naive customers ($s = 0$) assume at time 0 that they will not discount the sunk cost at all at time $\tau_0$ in making a decision. However, after the diagnosis, when the customer is about to make a decision at time $\tau_0$ about continuing with the service, she will, in fact, discount the sunk cost by $f\theta$. Thus, naive customers exhibit dynamically inconsistent preferences compared to sophisticated customers.
2.2 Sequence of events

The sequence of events under a two-part tariff for partially sophisticated customers is as follows (see Figure 1): (i) The service provider posts diagnosis fee $f$ and rate $r$ per unit time for treatment. (ii) Customers arrive for service if expected payment is anticipated to be less than expected value at time 0. Before diagnosis, the customer believes that she will continue with the service if her type is diagnosed to be $\tau \leq \bar{\tau}$ where $v_0 - fs\theta = r\bar{\tau}$, and will not otherwise. (iii) Next, after time $\tau_0$, the provider diagnoses a customer’s type as type $\tau$. (iv) Finally, customers now discount the sunk cost in deciding whether to continue with the service. Thus, a customer continues with the service if $\tau \leq \bar{\tau}$ and quits otherwise, where $v_0 - f\theta = r\tau$ (which is different from her initial belief $\bar{\tau}$ if $s < 1$).

3 Two-Part Tariff

In this section, we consider a monopoly firm that serves a single customer class comprised of boundedly rational customers who exhibit sunk cost bias ($\theta$) and are partially sophisticated ($s$). We analyze the provider’s optimal decisions and expected profit when using a two-part tariff. We also point out the outcomes in the special cases of rational customers ($\theta = 0$), as well as sophisticated ($s = 1$) and naive ($s = 0$) customers. We explore the impact of the level of sunk cost bias ($\theta$) on the optimal decision and profit. A two-part tariff is equivalent to charging two different rates during diagnosis and treatment because the diagnosis time is fixed. In particular, if we let $f = r_1\tau_0$ and $r = r_2$, the two-part tariff model is equivalent to a time-based contract with two different rates, $r_1$ and $r_2$, respectively, charged for diagnosis and treatment.
A representative customer arriving for service will (i) join the service at time 0 before diagnosis, and (ii) continue with the service at time \( \tau_0 \), upon learning of her type after diagnosis, only if the expected payment does not exceed the expected value from the service. First, consider the customer’s participation constraint at time 0, which requires that the expected payment does not exceed the expected value. The expected payment from a customer before diagnosis is given by:

\[
 r_1 \tau_0 + r_2 \int_0^{\bar{\tau}} \tau d\tau = r_1 \tau_0 + \frac{r_2}{2} \bar{\tau}^2.
\]

At time 0, the expected value from the service is \( v_0 \bar{\tau} \) because a customer believes at time 0 that the probability that she will continue with the service is \( \bar{\tau} \) and the value derived is \( v_0 \). Thus, a customer will start the service at time 0 only if:

\[
v_0 \bar{\tau} \geq r_1 \tau_0 + \frac{r_2}{2} \bar{\tau}^2.
\]

At time \( \tau_0 \), after the diagnosis, customers of type \( \tau \in [\tau, 1] \) will quit the service and, as a result, the service provider will only provide continued service to customers of type \( \tau \in [0, \tau] \). Hence, the service provider’s profit is given by:

\[
 (r_1 - c)\tau_0 + (r_2 - c) \int_0^{\bar{\tau}} \tau d\tau = (r_1 - c)\tau_0 + (r_2 - c) \bar{\tau}^2.
\]

The problem formulation is then:\footnote{In Appendix B, we consider a setting where value may increase over time, i.e., \( v(\tau) = v_0 + \kappa \tau \), where \( \kappa > 0 \). We show that the problem formulation can be transformed to this formulation with constant value. Moreover, our results in this section hold when \( \kappa < c \).}

\[
 \Pi_{tpt}^s(\theta) = \max_{r_1, r_2} \quad (r_1 - c)\tau_0 + (r_2 - c) \bar{\tau}^2 \quad \text{s.t.} \quad v_0 \bar{\tau} \geq r_1 \tau_0 + \frac{r_2}{2} \bar{\tau}^2, \quad \bar{\tau} = \min(\frac{v_0}{r_2} - \frac{r_1}{r_2} \tau_0 s\theta, 1), \quad \bar{\tau} = \min(\frac{v_0}{r_2} - \frac{r_1}{r_2} \tau_0 \theta, 1) \quad \bar{\tau} \geq 0, \bar{\tau} \geq 0.
\]

Let \( \Pi_{tpt}^s(\theta) \) denote the profit from the two-part tariff (tpt) when sophistication level is \( s \) and sunk cost bias is \( \theta \). As discussed previously, the first constraint captures the fact that the customer will join the service at time 0 only if the expected payment is less than or equal to the expected value. The second and third constraints help define the threshold values \( \bar{\tau} \) and \( \bar{\tau} \) respectively.

### 3.1 Rational Customers

We first consider the special case of rational customers who have no sunk cost bias – i.e., \( \theta = 0 \), which serves as a benchmark.
Proposition 1. When customers are rational ($\theta = 0$), the optimal two-part tariff is given by: (i) if $v_0 \geq c$, $\bar{\tau}^* = \tau^* = 1$, $r_1^* = \frac{\bar{v}_0}{2\bar{c}_0}$, $r_2^* = v_0$, $\Pi_{tpt} = v_0 - \frac{\bar{c}}{2} - c\tau_0$; (ii) if $v_0 < c$, $\bar{\tau}^* = \tau^* = \frac{\bar{v}_0}{c}$, $r_1^* = \frac{v_0^2}{2\bar{c}_0}$, $r_2^* = c$, $\Pi_{tpt} = \frac{v_0^2}{2c} - c\tau_0$.

When customer value $v_0$ is high relative to the cost ($c$) of serving them, and customers are rational, the pricing is such that all of the customers join and continue with the service. The fixed fee $f = v_0/2$ when $v_0 \geq c$ while it is less than $v_0/2$ otherwise. Moreover, the rate charged for treatment fully captures the value $v_0$ when $v_0 \geq c$, while it only covers the cost per unit $c$ when $v_0 < c$. Thus, when the value to cost ratio is lower, both diagnosis and treatment rates are lowered, fewer customers will obtain diagnosis and treatment, and the provider makes lower total profits. It can also be shown from the optimal values of $\bar{\tau}^*, \tau^*, r_1^*$, and $r_2^*$ that the payment for diagnosis ($=r_1^*\tau_0$) is equal to the expected payment for treatment ($=r_2^*\bar{\tau}^2/2$). Thus, as diagnosis time increases, the diagnosis rate is decreased but it does not impact the treatment or diagnosis fee.

3.2 Boundedly Rational Customers

We now consider the general case with boundedly rational customers. The following result characterizes the optimal two-part tariff that maximizes the profit $\Pi_{tpt}^*(\theta)$ when customers’ sophistication level is $s$ and sunk cost bias is $\theta$.

Proposition 2. (a) When customers are partially or fully naive, i.e. $s < 1$, (i) the optimal solution of the two-part tariff problem is: $\bar{\tau}^* = \min(\hat{\tau}, 1)$, $r_1^* = \frac{\bar{\tau}v_0}{2(1 - \theta s)}$, $r_2^* = v_0\left(\frac{1}{\hat{\tau}} - \frac{\theta s}{2 - \hat{\theta} s}\right)$, $\Pi_{tpt}^* = \frac{\tau(2 - \theta(s+1)^\theta)}{2(1 - \theta s)}$, where $\hat{\tau}$ is the value of $\tau$ that solves:

$$v_0\theta s \tau - \tau (1 - \theta s \tau) \frac{\partial}{\partial \tau} \left(1 - \theta s \tau / 2\right) = 0;$$

and $\hat{\tau} = \frac{(1 - \theta(s+1)^\theta/2)}{(1 - \theta s)}$ and $\frac{\partial}{\partial \tau} = \frac{2 - \theta(s+1)^\theta(2 - \theta s \tau)}{2(1 - \theta s \tau)^2}$; (ii) if $\frac{v_0}{c} \geq \frac{1 - \theta s}{\bar{\theta} s}$, then optimal $\bar{\tau}^* = 1$; (iii) the optimal profit increases with $\theta$ and decreases in $s$. (iv) $\bar{\tau}^*$ and $r_1^*$ increase in $\theta$ and $s$, and $r_2^*$ decreases in $\theta$ and $s$.

(b) When customers are sophisticated, i.e., $s = 1$, the optimal two-part tariff is given by: (i) if $v_0 \geq c$, $\bar{\tau}^* = \tau^* = 1$, $r_1^* = \frac{\bar{v}_0}{2\bar{c}_0 - \theta c}$, $r_2^* = v_0(\frac{2 - \theta}{2 - \theta s})$; (ii) if $v_0 < c$, $\bar{\tau}^* = \tau^* = \frac{\bar{v}_0}{c}$, $r_1^* = \frac{v_0^2}{2c\bar{c}_0(1 - \theta v_0/2c)}$, $r_2^* = c(\frac{2\theta - 2c/v_0}{s - 2c/v_0})$. Moreover, the optimal profit when customers are sophisticated ($s = 1$, for any $\theta$) is the same as for rational customers ($\theta = 0$).

Proposition 2 sheds light on the impact of bounded rationality on the optimal decisions and profits. First, we find that the optimal profit increases with sunk cost bias when $s < 1$. This is a surprising finding given that the presence of sunk cost bias suggests a lower willingness to pay
for treatment. From part (a) (iv) of Proposition 2, as the sunk cost bias increases, the provider increases the rate \( r_1 \) or equivalently the fixed fee \( f \) while reducing the rate \( r_2 \) for treatment. While an increase in \( f \) increases the sunk cost that the customer discounts after diagnosis, note that the customer will assume before diagnosis that she will discount it only partially (if at all), depending on the value of \( s \). The provider is able to exploit this partial (or full) naivete and charge a higher diagnosis fee. After diagnosis, the customer discounts the sunk cost fully and so the provider has to keep the rate low to ensure that customers continue with the service. This is why the rate for diagnosis increases and the rate for treatment decreases when sunk cost bias \( \theta \) increases. As the sunk cost bias increases, the provider can vary the values of \( r_1 \) (or \( f \)) and \( r_2 \) to take advantage of the customer’s naivete and increase profits.

We next consider the impact of sophistication level \( s \). When customers are fully sophisticated, the provider makes the same profit as when customers are rational and the sunk cost bias is irrelevant. When customers are more naive (\( s \) decreases), profits are higher and the provider makes the highest profit when customers are fully naive. Thus, we see that sunk cost bias itself may not impact profits if there is no dynamic inconsistency, as in the case of sophisticated customers. It is the dynamic inconsistency of customers in their belief about their sunk cost bias that allows the provider to extract greater profits.

### 3.3 Special Case: Conditional waiver of diagnosis fee

An interesting special case of the two-part tariff is commonly seen in practice in settings with a diagnosis phase – a fixed diagnosis fee is charged to all customers but it is waived if the customer continues with the treatment phase. For example, an appliance repair technician will charge a diagnosis fee, but it will be waived if the repair is performed by the technician. Such schemes are frequently used for computer and appliance repairs\(^5\). Since the diagnosis fee is waived for those continuing with the service, there is effectively no sunk cost and therefore no sunk cost bias. We refer to this scheme as a conditional TPT.

Recalling the definition of \( \tau \) in section 2.2 and noting that diagnosis fee is waived and there is no sunk cost, a customer at time \( \tau_0 \), being notified of type \( \tau \), will continue with the service if and only if \( \tau_0 \geq r_2 \tau \) and so \( \tau = \min(\frac{\tau_0}{r_2}, 1) \). Also, \( \tau = \min(\frac{\tau_0}{r_2}, 1) \) because the customer at time 0 also knows that the diagnosis fee will be waived if he continues with the service. Hence, the problem

formulation is as follows:

\[
\Pi_{ctpt} = \max_{r_1,r_2} \quad r_1 \tau_0 (1 - \tau) + (r_2 - c) \frac{\tau^2}{2} - c\tau_0 \\
\text{s.t.} \quad v_0 \bar{\tau} \geq r_1 \tau_0 (1 - \tau) + \frac{r_2 \tau^2}{2} \\
\bar{\tau} = \tau = \min(\frac{v_0}{r_2}, 1). 
\]

\(\Pi_{ctpt}\) is the profit of the conditional TPT scheme. The diagnosis fee \(r_1 \tau_0\) is multiplied by the term \((1 - \tau)\) in the objective function and in the first constraint because only customers who leave after diagnosis, which is given by the fraction \((1 - \tau)\), pay the diagnosis fee. Note that \(s\) and \(\theta\) do not appear in the formulation. We then have the following result on the optimal profit.

**Proposition 3.** The optimal profit of the conditional TPT is given by: if \(v_0 \geq c\), \(\tau = \bar{\tau} = 1\), \(r_0^* = v_0\), fixed fee is not charged to any customer since \(\tau = \bar{\tau} = 1\), \(\Pi_{ctpt} = v_0 - \frac{c}{2} - c\tau_0\); if \(v_0 < c\), \(\tau = \bar{\tau} = \frac{v_0}{c}\), \(r_0^* = c, r_1^* = \frac{v_0^2}{2\tau_0(c-v_0)}\), \(\Pi_{ctpt} = \frac{v_0^2}{2c} - c\tau_0\).

Note that the optimal profit is identical to that of the optimal two-part tariff with rational customers (see Proposition 1) in section 3.1. With a conditional waiver of the diagnosis fee, the sunk cost bias is irrelevant and so the provider can treat the customers as if they are rational. However, note that the optimal rates are different in Propositions 3 and 1. This is because the fixed diagnosis fee is waived only for customers leaving after diagnosis in the conditional TPT and so the rates are adjusted appropriately. When \(v_0 \geq c\), all customers are served and so the diagnosis fee is never charged and the rate charged for repair recovers the entire value.

### 4 Other Pricing Schemes

In this section, we consider various special cases of the two-part tariff that are commonly used in practice such as fixed fee, conditional fixed fee, single rate time-based fee and free diagnosis, and then compare them with each other as well as with the two-part tariff.

#### 4.1 Single-rate time-based fee

A single rate time-based fee is equivalent to a two-part tariff in which \(r_1 = r_2 = r\), i.e., rate \(r\) is identical in both phases of the service process and is commonly used in legal and consulting services, auto repair and several other industries (Lowendahl (2005)).\(^6\) The provider decides and posts the rate \(r\) before any diagnosis. The formulation for the single rate time-based scheme is similar to the two-part tariff formulation discussed in Section 3 except that \(r_1 = r_2 = r\), and is

\(^6\)See, for example: http://www.independent-consulting-bootcamp.com/consultant-rates.html
therefore relegated to the Appendix (in the proof of Proposition 4). Let \( \Pi_s(\theta) \) denote the profit from the time-based scheme \((t)\) when the sophistication level is \(s\) and sunk cost bias is \(\theta\). The optimal solution of the time-based scheme is provided in the next result.

**Proposition 4.** (i) Under the time-based scheme, if \( \frac{v_0}{\sqrt{k}} \geq c \) where \( k := 2\tau_0 + \theta^2 s^2 \tau_0^2 \), the optimal rate is \( r^* = \frac{v_0}{\sqrt{k}} \) and the provider’s profit is

\[
\Pi_s(\theta) = \left( \frac{v_0}{\sqrt{k}} - c \right) (\tau_0 + \frac{1}{2} \tau^2) = \left( \frac{v_0}{\sqrt{k}} - c \right) (\tau_0 + \frac{1}{2} (\sqrt{k} - \theta \tau_0)^2);
\]

otherwise the market is inactive. (ii) The provider’s profit \( \Pi_s(\theta) \) decreases in \( \theta \) and \( s \).

The optimal rate \( r^* \) is decreasing in \( s \). In particular, the optimal rates for naive \((s = 0)\) and sophisticated \((s = 1)\) customers are, respectively, \( \frac{v_0}{\sqrt{2}\tau_0} \) and \( \frac{v_0}{\sqrt{2\tau_0 + \theta^2 \tau_0^2}} \). Although intuitively, both naive and sophisticated customers discount sunk costs, naive customers do not recognize their bounded rationality at the time of joining the service. As a result, they mistakenly believe that they will not be deducting sunk costs and are effectively willing to pay more for the service than sophisticated customers at time 0. The provider exploits this fact by charging a higher rate and extracting more surplus from naive customers. It is interesting to note that the rate posted for rational consumers is the same as the rate for naive customers. This is because, in terms of the participation constraint (1) at time 0, naive customers essentially behave as if they are rational. In other words, they believe that they will not exhibit sunk cost bias similar to rational consumers, and the provider charges a rate consistent with this belief. On the other hand, sophisticated customers believe they will deduct sunk costs and so the provider has to correspondingly charge a lower rate to accommodate the sunk cost bias. The optimal rate \( r^* \) is also decreasing in sunk cost bias \( \theta \) (unless customers are naive) and in diagnosis time \( \tau_0 \). This is in contrast to the result under the two-part tariff wherein the optimal rate for diagnosis increases in \( \theta \) and \( s \), and the optimal rate for treatment decreases. When sunk cost bias or diagnosis time is high, the customer is less likely to join the service.

From part (ii) of the result, the provider makes higher profits if the market is comprised of naive \((s = 0)\) customers rather than sophisticated customers \((s = 1)\). Both decreasing levels of sophistication and increasing values of the sunk cost bias are indicators of greater bounded rationality. While lower levels of sophistication allow the provider to extract greater profits, increased sunk cost bias \((\theta)\) results in lower profits. Greater naivete indicates an inconsistency in the willingness to pay before and after diagnosis, which can be exploited by the provider. The lower profit when \( \theta \) increases is in contrast to the result under the two-part tariff, in which the profit increases for higher \( \theta \). This is because the two-part tariff has two different rates for diagnosis and treatment that can be used to exploit customers’ inconsistent beliefs about their sunk cost bias at time 0 and
When a provider is restricted to a single rate throughout the diagnosis and treatment phases under the time-based scheme, he cannot exploit a customer’s uneven perception of the sunk cost effect at different stages of the service process.

### 4.2 Fixed Fee and Conditional Fixed Fee

Under a fixed fee, the provider charges a fee $f$ and serves all customers. A fixed fee, more commonly called a flat fee, is used in many industries. The provider’s profit is $f - c(\tau_0 + \int_0^1 \tau d\tau)$ and a customer will join the service if $v_0 \geq f$. It then follows that the monopolist provider’s profit under the fixed fee is $\Pi_f = v_0 - c(\tau_0 + \frac{1}{2})$ and the optimal fee charged is $f = v_0$. Optimal decisions and profits are not impacted by bounded rationality with a fixed fee. A drawback of the fixed fee is that the provider is forced to serve customers with high service needs who are not profitable. A variant of the fixed fee scheme that does not suffer from this drawback is a conditional fixed fee. In this scheme, all customers are charged the same fixed fee $f$ upon joining the service. However, the provider may reject some customers after diagnosis and such customers are refunded the fee $f$.

For example, a lawyer may quote a fixed fee for certain services but then offer it only for routine cases. The provider will naturally reject those customers whose treatment may take too long, i.e. customers of type $\tau > \bar{\tau}$ upon diagnosis, where $\bar{\tau}$ is a threshold value. The conditional fixed fee, similar to the simple fixed fee, does not depend on $\theta$ or $s$. We then have the following result.

**Proposition 5.** The provider’s profit using the conditional fixed fee scheme is: if $v_0 < c$, $\Pi_{cf} = \frac{v_0^2}{2c} - cv_0$ and $\Pi_{cf} - \Pi_f = \frac{(v_0 - c)^2}{2c}$; if $v_0 \geq c$, $\Pi_{cf} = v_0 - \frac{c}{2} - cv_0 = \Pi_f$. When $\theta = 0$ or $s = 0$, profit is equal to that from a two-part tariff.

The conditional fixed fee scheme generates a higher profit than a simple fixed fee scheme when $v_0 < c$, and it can achieve the same profit as the two-part tariff when customers are rational or fully sophisticated. However, it has a potential drawback: customers learn only after diagnosis if they will obtain treatment. They have no a priori knowledge about what their type is going to be and this causes uncertainty. Moreover, this pricing scheme may appear to be unfair to a potential customer as only the provider has discretion over who they will treat. Therefore, rejected customers may leave the service dissatisfied, even with a free diagnosis. This drawback may adversely impact the perception of the service and the probability that a potential customer will even arrive to obtain a diagnosis.

---

7See, for example: https://its.unl.edu/helpcenter/repairs-rates for computer repair and http://www.123triad.com for website design services. The site fixedpricetrade.com offers numerous home related services at fixed prices.
4.3 Free diagnosis

We next consider a special case of the two-part tariff wherein the provider charges zero fees for diagnosis (i.e., \( r_1 = f = 0 \)) and charges a time-based fee for the treatment phase.\(^8\) Free diagnosis is also offered in practice—for example, in auto repair shops, computer repair, etc.\(^9\). Not charging a diagnosis fee may be attractive in our setting because boundedly rational consumers discount the (sunk) diagnosis costs in deciding whether to get treatment and the sunk cost is zero in this case. Let \( \Pi_{fd} \) denote the profit under free diagnosis.

**Proposition 6.** When \( v_0 \geq 2c \), the service provider charges \( r_2 = v_0 \) and makes a profit of \( \Pi_{fd} = \frac{1}{2}v_0 - c\tau_0 + \frac{1}{2}c \), which is lower than the profit under a fixed fee; When \( v_0 \leq 2c \), the service provider charges \( r_2 = 2c \) and makes a profit of \( \Pi_{fd} = \frac{1}{8}v_0^2 - c\tau_0 \).

When \( v_0/c \geq 2 \) (the value is much greater than the cost of effort per unit time), the provider is better off using a fixed fee. We will see later that a free diagnosis may not be better than fixed fee even when \( v_0/c < 2 \). In this scenario, the provider has to charge a high rate, \( r_2 = 2c \), during treatment to compensate for the fact that the diagnosis was free. This turns out to be potentially problematic as some of the high type customers will walk away after diagnosis. This is why the profit is \( \frac{1}{8}v_0^2 - c\tau_0 \), which is less than the profit from conditional fixed fee \( \frac{1}{2}v_0^2 - c\tau_0 \). Another potential drawback of a free diagnosis is that customers find out their type only after diagnosis and may be concerned about potentially high treatment costs. While we have captured this in an expected sense in our model, a free diagnosis may generate some uncertainty and skepticism about treatment costs prior to diagnosis. Of course, a customer is free to leave after the free diagnosis. Since there is no diagnosis fee, sunk cost bias does not impact optimal decisions and profits.

4.4 Comparison of Pricing Schemes

We now compare profits across these performance schemes and explore how bounded rationality \( s \) and \( \theta \) may impact their relative performance. We begin by noting that a two-part tariff dominates all the specialized pricing schemes in terms of profit. We next present a result that compares the performance of the four pricing schemes discussed previously (in sections 3.3, 4.1, 4.2, 4.3).

**Proposition 7.** (a) \( \Pi_{ctpt} = \Pi_{cf} \geq \Pi_f; \Pi_{cf} \geq \Pi_{fd}; \Pi_{cf} \geq \Pi_{st}(\theta) \) for all \( s, \theta \); (b) If \( v_0 \geq 0.536c \), then \( \Pi_f \geq \Pi_{fd} \). (c) Let \( L(s, \tau_0) \) and \( U(s, \tau_0) \)\(^{10}\) represent lower and upper bounds respectively on the value of \( \frac{v_0}{c} \). Then: (i) If \( L(s, \tau_0) \leq \frac{v_0}{c} \leq U(s, \tau_0) \), there exists \( \tilde{\theta}_s \in [0, 1] \) such that \( \Pi_{st}(\theta) \geq \Pi_f \) if

---

\(^8\)The case where \( r_2 = 0 \) and \( r_1 \) (or \( f \)) > 0 is equivalent to a simple fixed fee and is therefore omitted.


\(^{10}\)The formulas for \( L(s, \tau_0) \) and \( U(s, \tau_0) \) are provided in the proof.
θ ≤ \tilde{\theta}_s \text{ and } \Pi^*_t(\theta) < \Pi_f \text{ if } \theta > \tilde{\theta}_s, \text{ where } \tilde{\theta}_s \text{ decreases with } s. \ (ii) If \ \frac{v_0}{c} \geq U(s, \tau_0) \text{ fixed fee payment is the dominant strategy; (iii) If } \frac{v_0}{c} < L(s, \tau_0),\text{ time-based payment is the dominant strategy.}

The conditional fixed fee dominates the fixed fee and the single rate time-based fee. But it may be perceived as unfair since the provider maintains control over who obtains service and the customer finds this out only after diagnosis. The conditional TPT achieves the same performance as the conditional fixed fee but does not suffer from the drawback of turning away customers involuntarily. This is perhaps why we see it frequently used in practice.

The free diagnosis scheme is always dominated by the conditional fixed fee and is dominated by the fixed fee or the time-based fee in most scenarios. The simple fixed fee outperforms free diagnosis when \( v_0 \geq 0.536c \). Recall that the fixed fee scheme is profitable only when \( v_0 > 0.5c + c\tau_0 \). In turn, this implies that the fixed fee scheme dominates free diagnosis in almost the entire range where the fixed fee is profitable. The range in which free diagnosis is profitable and better than the fixed is very narrow. This is clear from the fact that \( \Pi_{fd} = \frac{1}{8}v^2/c - c\tau_0 \) and so \( \Pi_{fd} \geq 0 \) and \( v_0 \leq 0.536c \) can be true only if \( \tau_0 \leq \frac{0.536^2}{8} = 0.0359 \), i.e. if the diagnosis time is very small and value from the service is much smaller than the cost of serving customers. It is a simple scheme and appears to be attractive because consumers like free diagnosis. At first glance, it also appears to be attractive in an environment with a sunk cost bias, since free diagnosis eliminates the issue of sunk cost bias. However, while providers may be tempted to offer free diagnosis, it is largely dominated by the other schemes. In particular, it is better to offer the conditional TPT which offers free diagnosis only if the customer continues with the service.

Next, we discuss part (c) and compare the single-rate time-based fee and fixed fee which are commonly used in practice. When the value from the service \( v_0 \) is small and the service cost per unit time \( c \) is high, it is costly to serve high type customers and the time-based scheme displays its virtue by screening out customers with undesirably large \( \tau \). The fixed fee scheme cannot screen customers and hence will be a dominant strategy only when the ratio \( v_0/c \) is high. This is true whether consumers are rational (\( \theta = 0 \)) or not. It is when \( \frac{v_0}{c} \) is in a medium range, as defined in Proposition 7, that the sunk cost bias \( \theta \) and the sophistication level \( s \) have an impact on the pricing scheme. A time-based contract is more attractive than a fixed fee when the sunk cost bias (\( \theta \)) is small. When \( \theta \) increases, this effectively lowers the amount the customer is willing to pay for treatment, which makes the time-based scheme less attractive, but the fixed fee is not impacted. The threshold level \( \tilde{\theta}_s \) beyond which the fixed fee becomes attractive decreases with the sophistication level. Proposition 7 suggests that the provider can extract more surplus from less sophisticated customers for any given \( \theta \). Thus, the range of \( \theta \) values, wherein the time-based scheme dominates the fixed scheme, is broader when customers are more naive (smaller \( s \)). This is because, although the profit from serving any customer decreases with \( \theta \), it decreases at a greater
Table 1: Summary of Different Payment Schemes and the Profit Changes with $\theta$ and $s$

<table>
<thead>
<tr>
<th>Contract Type</th>
<th>Diagnosis Fee</th>
<th>Treatment Fee for Type $\tau$</th>
<th>Profit Changes with $\theta$</th>
<th>Profit Changes with $s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-Part Tariff</td>
<td>$r_1\tau_0$</td>
<td>$r_2\tau$</td>
<td>Increase</td>
<td>Decrease</td>
</tr>
<tr>
<td>Time-Based Fee</td>
<td>$r\tau_0$</td>
<td>$r\tau$</td>
<td>Decrease</td>
<td>Decrease</td>
</tr>
<tr>
<td>Fixed Fee</td>
<td>$f$</td>
<td>0</td>
<td>Does not change</td>
<td></td>
</tr>
<tr>
<td>Conditional Fixed Fee</td>
<td>$f$</td>
<td>Reject with refund of $f$ if $\tau \geq \bar{\tau}$</td>
<td>Does not change</td>
<td></td>
</tr>
<tr>
<td>Conditional Two-Part</td>
<td>$r_1\tau_0$</td>
<td>$r_2\tau$ if $\tau \leq \bar{\tau}$</td>
<td>Does not change</td>
<td></td>
</tr>
<tr>
<td>Part Tariff</td>
<td></td>
<td>Refund diagnosis fee if $\tau &gt; \bar{\tau}$</td>
<td>Does not change</td>
<td></td>
</tr>
<tr>
<td>Free Diagnosis</td>
<td>0</td>
<td>$r_2\tau$</td>
<td>Does not change</td>
<td></td>
</tr>
</tbody>
</table>

rate when customers are more sophisticated.

Table 1 contains a summary of the various pricing schemes along with the impact of bounded rationality on their profit performance. The fixed fee, conditional fixed fee, conditional TPT and free diagnosis are not impacted by bounded rationality. Hence, these pricing schemes have the advantage that the provider does not have to gather information on the values of $s$ and $\theta$ in the population. On the other hand, the optimal rate in the time-based fee and two-part tariff do depend on these values and allow the provider to exploit bounded rationality to achieve higher profits.

The two-part tariff has the advantage of achieving the highest profit and is seemingly fair because customers decide whether to continue with the service after diagnosis. This also implies that it would be a mistake for the firm to assume that consumers are rational and ignore bounded rationality unless $v_0/c$ is high (significantly higher than 1). If consumers are boundedly rational, the firm has to make an effort to understand the degree of sophistication and the level of sunk cost bias in determining the optimal rate(s) to charge in the time-based fee or two-part tariff. For instance, suppose the firm were to assume consumers are rational – i.e., $\theta = 0$ (even though they are not) – and price accordingly. It can be easily shown in this case that the rate charged in the time-based fee would be such that the participation constraint $v_0\bar{\tau} \geq r(\tau_0 + \frac{1}{2}\tau^2)$ would be violated and therefore the firm will make zero profits.\(^\text{11}\) Moreover, the two-part tariff can actually exploit naive customers and achieve higher profits than when customers are rational and in fact, profits are higher when sunk cost bias is greater. Hence, it is important that a provider understand whether its consumers exhibit bounded rationality and estimate the level of sophistication and extent of sunk cost bias in its customer base. The conditional fixed fee and conditional TPT are good alternatives

\(^{11}\)Note that the only exception is when customers are naive ($s = 0$). In this case, the service provider does not lose the market although he makes less profit than in the scenario in which he prices knowing that consumers are naive.
if the provider does not want to estimate the bounded rationality parameters but still wants a highly profitable scheme.

### 4.4.1 Numerical Study

We performed a numerical study to better understand the relative performance of the different pricing schemes. While Proposition 7 provides analytical results on the relative performance of the different schemes, it does not shed light on the extent to which one pricing scheme may outperform another, or how the relative performance varies with the level of bounded rationality. We considered the following parameters in the study. We set $c = 10$ and varied $v_0$ from 5 to 15 in steps of 1 as profitability depends on the ratio $v_0/c$. This allowed us to explore the performance when $v_0 > c$, as well as when $v_0 < c$. The other parameters were chosen as follows: $\tau_0 \in [0.1, 0.2, 0.3, 0.4, 0.4]$, $\theta \in [0.0, 0.2, 0.4, 0.6, 0.8]$, and $s \in [0.0, 0.2, 0.4, 0.6, 0.8, 1.0]$. These parameter ranges allowed us to test performance across a wide spectrum of problem instances. We did not consider $\tau_0$ values greater than 0.5 because it is unlikely that diagnosis time will comprise more than 50% of the total time in realistic scenarios. Thus, we arrived at 1650 possible problem instances. We tested the performance of the four pricing schemes as well as the two-part tariff. Out of these 1650 instances, we restricted our attention to those for which the two-part tariff (which has the highest profit) is profitable – 1244 problem instances.

Table 2 shows the performance of four pricing schemes relative to the two-part tariff for various quartiles across the 1244 problem instances. We did not include free diagnosis because it performs very poorly – in fact, in most instances, it had negative profits (and would therefore not be adopted) while the other schemes had positive profits. For each of the pricing schemes, the numbers in the table indicate the profit of these schemes as a % of two-part tariff (TPT to be concise) profit – i.e., $100 \times \left( \frac{\text{Profit from Simple Pricing Scheme}}{\text{Profit from TPT}} \right)$. For instance, the profit of a time-based scheme is equal to 62.5% of that of TPT on average across the 1244 problem instances. At the

<table>
<thead>
<tr>
<th>Simple Pricing Scheme</th>
<th>Time-based</th>
<th>Conditional TPT/Fixed Fee</th>
<th>Fixed Fee</th>
</tr>
</thead>
<tbody>
<tr>
<td>25th percentile</td>
<td>50.9%</td>
<td>91.7%</td>
<td>80.6%</td>
</tr>
<tr>
<td>Median</td>
<td>64.1%</td>
<td>100%</td>
<td>98.2%</td>
</tr>
<tr>
<td>75th percentile</td>
<td>50.9%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Maximum</td>
<td>100.0%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Average</td>
<td>62.5%</td>
<td>86.4%</td>
<td>61.8%</td>
</tr>
</tbody>
</table>

Table 2: Performance of the Other Pricing Schemes Relative to Two-Part Tariff
25th percentile of the 1244 problem instances, the profit of the conditional TPT (or conditional fixed fee) is 91.7% of the profit under TPT. It is clear from the table that the conditional TPT (equivalently conditional fixed fee) is a good alternative to TPT and does as well as TPT in a majority of the problem instances. However, there are a few problem instances where they can do poorly – specifically, these are cases where $\theta$ is high and $v_0 < c$. While the time-based fee shows somewhat better performance than the fixed fee on average, the fixed fee does better in more problem instances. Again, fixed fee does poorly when $\theta$ and $\tau_0$ are high and $v_0 < c$.

Figures 2(a) and 2(b) illustrate the extent to which TPT outperforms the other three pricing schemes for different values of $\theta$ and $s$. All three schemes do worse relatively as $\theta$ increases but the time-based scheme and fixed fee do much worse. Thus, as sunk cost bias increases, it is best to use TPT and conditional fixed fee or conditional TPT is the next best alternative. The impact of sophistication level $s$ is quite different. The outperformance of TPT relative to the time-based scheme is insensitive to $s$. The outperformance of TPT as compared to fixed fee and conditional fixed fee is a bit more sensitive to $s$ but not monotonic in $s$.

5 Model with Heterogeneous Sunk Cost Bias

In the previous two sections, we assumed a scenario wherein customers are homogeneous with respect to their sunk cost bias and sophistication level. In reality, a market may be comprised of customers with different levels of bounded rationality. Next, we first consider a setting where there are two classes of customers with different levels of sunk cost bias and explore the type(s) of two-part tariff contracts a provider may offer in such a setting. In addition, we also explore
the loss in profit the service provider will experience by mistakenly treating customers as if they are homogeneous when in fact they are not. We then provide a similar analysis for the case in which there are two classes of customers with different sophistication levels. Finally, we consider the setting in which customers are heterogeneous in both dimensions of bounded rationality.

5.1 Heterogeneity in Sunk Cost $\theta$

There is one customer class with sunk cost bias $\theta_H$ and another with sunk cost bias $\theta_L$, and the proportions of the two classes are, respectively, $\lambda$ and $(1 - \lambda)$. We assume that the provider knows the proportion $\lambda$ but does not know if a customer is of type $\theta_H$ or $\theta_L$ before providing the service. As in Section 3, the provider offers a two-part tariff with rates $r_1$ and $r_2$. The problem formulation in this heterogeneous case is as follows:

$$\max_{r_1, r_2} \lambda((r_1 - c)\tau_0 + (r_2 - c)\frac{1}{2}\tau_2^2_H) + (1 - \lambda)((r_1 - c)\tau_0 + (r_2 - c)\frac{1}{2}\tau_2^2_L)$$

s.t.   $v_0\bar{\tau}_H \geq r_1\tau_0 + \frac{r_2}{2}\bar{\tau}_H^2$  (9)

$$v_0\bar{\tau}_L \geq r_1\tau_0 + \frac{r_2}{2}\bar{\tau}_L^2$$  (10)

$$\bar{\tau}_L = \min\left(\frac{v_0}{r_2} - \frac{r_1}{r_2}\theta_Ls\tau_0, 1\right)$$  (11)

$$\bar{\tau}_H = \min\left(\frac{v_0}{r_2} - \frac{r_1}{r_2}\theta_Hs\tau_0, 1\right)$$  (12)

$$\tau_L = \min\left(\frac{v_0}{r_2} - \frac{r_1}{r_2}\theta_Ls\tau_0, 1\right)$$  (13)

$$\tau_H = \min\left(\frac{v_0}{r_2} - \frac{r_1}{r_2}\theta_Hs\tau_0, 1\right).$$  (14)

The first (second) constraint ensures that customers with high (low) sunk cost bias will join the service at time 0 only if the expected payment is less than or equal to the expected value from the service. Constraints (9) to (12) define the threshold values $\bar{\tau}_L$, $\bar{\tau}_L$, $\bar{\tau}_H$, and $\bar{\tau}_H$ respectively, and are similar to the threshold definition constraints in the formulation in Section 3. We then have the following result that shows that the service provider cannot be better off by offering two different two-part tariff contracts to customers with heterogeneous sunk cost bias $\theta_L$ and $\theta_H$. This is because customers with low sunk cost bias $\theta_L$ would always prefer the contract provided to those with high sunk cost bias $\theta_H$ when the service provider offers two different contracts. Hence, the service provider instead offers only one two-part tariff contract.

**Proposition 8.** Let $(r^*_{1H}, r^*_{2H})$ and $(r^*_{1L}, r^*_{2L})$ be the optimal two part tariff to offer when serving only $\theta_H$ type and only $\theta_L$ type customers (i.e. they are the optimal solution of the problem (1)-(5) when $\theta = \theta_H$ and $\theta = \theta_L$), respectively. The optimal solution of the above optimization problem is
Figure 3: Profit Loss from Treating Customers with Heterogeneous Sunk Cost Bias as Homogeneous $(\tau_0 = 0.3, c = 5, v_0 = 4.5, s = 0.5, \theta_H = 0.6, \theta_L = 0.2)$

then to either post $(r^*_1, r^*_2)$ and serve both types or post $(r^*_1, r^*_2)$ and serve only the low type.

Thus, the provider either serves both types or serves only the low type and will choose the option that results in higher profits. Serving only the low type is likely to be more profitable when $\lambda$ is small. Interestingly, the optimal two-part tariff the provider would offer when serving both types is the same as what he would offer if he was only serving the high type. We next consider the profit loss of the service provider if he mistakenly treats the two classes of customers as if they are homogeneous with a weighted average sunk cost bias.

**Proposition 9.** If the service provider treats customers with heterogeneous sunk cost bias as homogeneous with type $\theta = \lambda \theta_H + (1 - \lambda) \theta_L$ and offers a two-part tariff contract $(r^*_1(\theta), r^*_2(\theta))$, the customer with sunk cost bias $\theta_H$ will then not join the service and the customer with sunk cost bias $\theta_L$ will have a positive surplus.

Thus, we see that the provider may be worse off in two respects by treating the heterogeneous customers as homogeneous: the high type will not join the service and the provider will have to give a positive surplus to the low type and hence achieve lower profits. We now compare the optimal profits of the two strategies, viz., treating customers as heterogeneous and homogeneous. Figure 3 shows how the provider’s profit changes with $\lambda$ in the two scenarios. The yellow curve represents the service provider’s optimal profit when considering the customers as belonging to two classes with different sunk cost bias values. The blue curve represents the service provider’s profit when treating customers as if they are homogeneous with sunk cost bias $\theta = \lambda \theta_H + (1 - \lambda) \theta_L$. When $\lambda$ (the proportion of high type customers) is small, it is optimal for the service provider to serve only $\theta_L$ type customers. In this case, the optimal profit in the heterogeneous case only slightly dominates the profit made by the service provider when treating heterogeneous customers as homogeneous.
However, when $\lambda$ becomes large, it becomes optimal for the service provider to serve both types of customers. The difference between the optimal profit in the heterogeneous case and the profit when treating customers as homogeneous increases because the service provider cannot attract the high type customers in the latter case and has to forgo the profits from this customer class.

5.2 Heterogeneity in Sophistication Level $s$

We now consider the setting with two different sophistication levels $s$. We assume that a customer has sophistication level $s_H$ with probability $\gamma$ and sophistication level $s_L$ with probability $1 - \gamma$. As in the previous section, the optimal strategy could be to serve both types or serve only the low type and the problem formulation is different in each case. Note that, different from the analysis in Section 5.2, the heterogeneous sophistication level only affects customers’ expectation to quit ex-ante but not the actual participation ex-post. That is, different $s_H$ and $s_L$ levels lead to different $\bar{\tau}_H$ and $\bar{\tau}_L$, but $\tau$ is only affected by $\theta$ and thus remains the same.

We then have the following result characterizing the optimal solution (the problem formulation for the two cases are included in the proof of Proposition 10).

Proposition 10. Let $(r_{1H}^*, r_{2H}^*)$ be the optimal two-part tariff offered when serving only $s_H$ type customers and $(r_{1L}^*, r_{2L}^*)$ when serving only $s_L$ type customers. The optimal solution of the above optimization problem is then either to post $(r_{1H}^*, r_{2H}^*)$ and serve both types or post $(r_{1L}^*, r_{2L}^*)$ and serve only the low type.

Next, as in the previous subsection, we identify the profit loss suffered by the service provider if he mistakenly treats the two classes of customers as if they are homogeneous with a weighted average sophistication level.

Proposition 11. If the service provider treats customers with heterogeneous sophistication levels as homogeneous with type $s = \gamma s_H + (1 - \gamma) s_L$, and offers a two-part tariff contract $(r_1^*(s), r_2^*(s))$, then the customer with sophistication level $s_H$ will not join the service and the customer with sophistication level $s_L$ will have positive surplus.

We now compare the optimal profit of the two strategies, and in Figure 4 we show how the provider’s profit changes with $\gamma$ in the two scenarios. The yellow curve represents the service provider’s optimal profit when considering the customers as heterogeneous with different sophistication levels. The blue curve represents the service provider’s profit when treating customers as if they are homogeneous with sophistication level $s = \gamma s_H + (1 - \gamma) s_L$. As in Figure 3, when $\gamma$ is small, it is optimal for the service provider to serve only $s_L$ type customers in the heterogeneous case. In this case, the optimal profit only slightly dominates the profit made by the service provider.
when treating heterogeneous customers as homogeneous. However, when \( \gamma \) becomes large, it becomes optimal for the service provider to serve both types of customers. The difference between the optimal profit and the service provider’s profit when treating customers as homogeneous increases because the service provider cannot attract high type customers in the latter case.

### 5.3 Heterogeneity in both dimensions

We now consider the scenario in which customers are heterogeneous in both sunk cost bias and sophistication level. Specifically, a customer has sunk cost bias \( \theta_H \) with probability \( \lambda \) and \( \theta_L \) with probability \( 1 - \lambda \). In addition, a customer’s sophistication level is \( s_H \) with probability \( \gamma \) and \( s_L \) with probability \( 1 - \gamma \). Thus, there are four classes or types of customers with different levels of sunk cost bias and sophistication level: \((\theta_H,s_H),(\theta_H,s_L),(\theta_L,s_H),(\theta_L,s_L)\). The formulation is similar to the ones provided in the previous two sub-sections (but more complex) and is therefore omitted. The participation constraint for the type \((\theta_i,s_j),i,j=H\) or \(L\), to join the service is \( v_0 \tau_{ij} \geq r_1 \tau_0 + \frac{r_2}{r_2} \tau_{ij}^2 \), where \( \tau_{ij} = \min(\frac{r_2}{r_2} - \frac{r_1}{r_2} \theta_i s_j \tau_0, 1) \). Note that the participation constraint, which will eventually determine the optimal two-part tariff, is only related to the term \( \theta_i s_j \).

Thus, the optimal two-part tariff contract that a provider would post is only related to the product of \( \theta_i \) and \( s_j \), \( i,j=H \) or \( L \), and so the optimal two-part tariff contract the provider should offer is one among those in: \((r_1^*(\theta_H s_H),r_2^*(\theta_H s_H)),(r_1^*(\theta_H s_L),r_2^*(\theta_H s_L)),(r_1^*(\theta_L s_H),r_2^*(\theta_L s_H))\), and \((r_1^*(\theta_L s_L),r_2^*(\theta_L s_L))\), each of which corresponds, respectively, to the optimal two-part tariff contract when facing only \((\theta_H,s_H),(\theta_H,s_L),(\theta_L,s_H),(\theta_L,s_L)\) type customers. Since \( \theta_H s_H \geq \max(\theta_H s_L,\theta_L s_H) \geq \min(\theta_H s_L,\theta_L s_H) \geq \theta_L s_L \), the service provider attracts only \((\theta_L,s_L)\) type customers by offering \((r_1^*(\theta_L s_L),r_2^*(\theta_L s_L))\). At the other extreme, the service provider attracts all four classes of customers by offering \((r_1^*(\theta_H s_H),r_2^*(\theta_H s_H))\). Thus several possible scenarios exist.
Table 3: The Optimal Two-Part Tariff Contract to Offer to Customers with Heterogeneous Sunk Cost Bias and Sophistication Level

<table>
<thead>
<tr>
<th>Contract/Customer Type</th>
<th>$\theta_{HSL} \geq \theta_{LSH}$</th>
<th>$\theta_{HSL} \leq \theta_{LSH}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\theta_{L,SL}$</td>
<td>$\theta_{H,SL}$</td>
</tr>
<tr>
<td>$(r_1^<em>(\theta_{L,SL}), r_2^</em>(\theta_{L,SL}))$</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>$(r_1^<em>(\theta_{H,SL}), r_2^</em>(\theta_{H,SL}))$</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$(r_1^<em>(\theta_{L,SH}), r_2^</em>(\theta_{L,SH}))$</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>$(r_1^<em>(\theta_{H,SH}), r_2^</em>(\theta_{H,SH}))$</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

depending on the parameters. The complete characterization of the contracts that the provider could offer, and the customer types who will join the service, are provided in Table 3. The optimal two-part tariff that a provider would offer, among the four alternatives identified in the Table, will depend on the profits from each alternative, which in turn will depend upon the relative proportions of each customer class and other problem parameters.

We now consider the scenario in which the service provider mistakenly treats customers with a heterogeneous sunk cost bias and sophistication level as homogeneous. Let $\xi = \lambda \gamma \theta_{H,H} + \lambda (1 - \gamma) \theta_{H,L} + (1 - \lambda) (1 - \gamma) \theta_{L,L}$. In this case, the service provider will offer a single two-part tariff contract $(r_1^*(\xi), r_2^*(\xi))$ to all the customers. Since $\theta_{L,SL} \leq \xi \leq \theta_{H,SH}$, the service provider will not be able to attract $(\theta_{H,SH})$ type customers as was true in the previous two settings, and will leave a positive surplus to $(\theta_{L,SL})$ type customers. In addition, the service provider will attract $(\theta_{H,SL})$ type customers only if $\theta_{H,SL} \leq \xi$ (and similarly attract $(\theta_{L,SH})$ type customers only if $\theta_{L,SH} \leq \xi$) and again leave them with a positive surplus. The service provider will not attract $(\theta_{H,SL})$ type and $(\theta_{L,SH})$ type customers if $\theta_{H,SL} > \xi$ and $\theta_{L,SH} > \xi$ respectively. It is then straightforward to compare the optimal profit using the two strategies as in the previous two subsections. Clearly, the provider will be worse off when the proportion of high types is higher along either or both dimensions, because he will not be able to attract some of the customer classes when treating all of them as one class and he will also have to give surplus to some types. Thus, the provider may be able to treat the customer classes as homogeneous and suffer minimal profit loss only when the proportion of low type is large along both dimensions of bounded rationality.

6 Conclusion

The sunk cost bias is a commonly observed phenomenon. We investigate the impact of the sunk cost bias on the relative attractiveness of several commonly used pricing schemes by a monopolist provider in a service setting with diagnosis and treatment phases. Further, we consider customers with varying levels of dynamic inconsistency in their sunk cost bias before and after diagnosis. The
two-part tariff scheme, due to its flexibility in charging different rates for diagnosis and treatment, is effective in exploiting a customer’s bounded rationality. In fact, it achieves higher profits when consumers have greater sunk cost bias and when they are more naive. The conditional fixed fee also does well despite charging a fixed fee because it has the flexibility to deny treatment to customers requiring high treatment time. Even the simple fixed fee and the time-based fee do well despite their simplicity and lack of flexibility; the dominance of one over the other depends on the levels of naivete and sunk cost bias and the ratio of value derived from service to the cost of providing service. A free diagnosis together with a time-based treatment fee sounds attractive but performs poorly relative to the other pricing schemes. The analysis and insights can be partially extended to a duopoly setting with the providers restricted to using a fixed fee or single rate time-based fee (details available from the authors).

We also consider a scenario wherein the market consists of two customer classes with different levels of sunk cost bias or different levels of sophistication. We find that it is not more profitable for the provider to offer two different two-part tariff contracts to serve both segments in either scenario. When customers are heterogeneous in both sunk cost bias and sophistication level, the provider chooses one out of four possible two-part tariff contracts that will attract one or more of the four segments. We find that it is important for the provider to recognize the heterogeneity in the customer base unless the proportion of customers with high sunk cost bias and sophistication level is very low.

Overall, we find that both the sunk cost bias and the degree of dynamic inconsistency can have a substantial impact on the choice of pricing scheme in diagnostic services. This is especially true when the ratio of value from the service to the cost of providing service \( (v_0/c) \) is low. Hence, it is important for a service provider to carefully understand the level of bounded rationality in its customer base in both dimensions when choosing a pricing scheme.

**Acknowledgments**

We are grateful to the anonymous review team (Senior Editor and reviewers) for their valuable comments and suggestions that have improved the paper along a number of dimensions, including the modeling, analysis, and exposition and for suggesting some of the pricing schemes studied.

The authors made equal contributions to this work. Professor Tong’s work was supported by the National Science Foundation of China under No.71772137 and he is the corresponding author.
References


Online Appendix to “Pricing Diagnosis-Based Services When Customers Exhibit Sunk Cost Bias”

Appendix A: Proofs

This Appendix contains proofs of the Propositions and Lemmas.

Proof of Proposition 1. When customers are rational, $\theta = 0$ and $\bar{\tau} = \tau$. Then the service provider’s profit under two-part tariff is

$$
\Pi_{\text{tpt}}^\theta(0) = \max_{r_1, r_2} (r_1 - c)\tau_0 + (r_2 - c)\frac{\tau^2}{2}
$$

s.t. \hspace{1cm} v_0\bar{\tau} \geq r_1\tau_0 + \frac{r_2}{2}\bar{\tau}^2 

(15)

$$
\bar{\tau} = \tau = \min\left\{ \frac{v_0}{r_2}, 1 \right\} 

(16)

\bar{\tau} \geq 0, \tau \geq 0.
$$

The constraint (15) is binding at optimality. Thus, the service provider’s profit when $\theta = 0$ is

$$
\Pi_{\text{tpt}}^\theta(0) = v_0\bar{\tau} - c\tau_0 - \frac{c}{2}\bar{\tau}^2.
$$

Taking the derivative of the objective function $\Pi_{\text{tpt}}^\theta(0)$ with respect to $\bar{\tau}$, we have:

$$
\frac{d\Pi_{\text{tpt}}^\theta(0)}{d\bar{\tau}} = v_0 \left( 1 - \frac{c\bar{\tau}}{v_0} \right).
$$

(i) If $v_0 \geq c$, then $\frac{d\Pi}{d\bar{\tau}} \geq 0$ for $\bar{\tau} \in [0, 1]$ and so $\bar{\tau}^* = 1$. As a result, $r_2 = v_0$ and $r_1 = \frac{v_0}{2\tau_0}$. (ii) If $v_0 < c$, then $\frac{d\Pi}{d\bar{\tau}} = 0$ when $\bar{\tau} = \frac{v_0}{c}$ and so $\bar{\tau}^* = \frac{v_0}{c}$. Hence, $r_2 = c$ and $r_1 = \frac{v_0}{2c\tau_0}$. ■

Proof of Proposition 2. (a) We first consider the case $s < 1$.

(i) We first show that the constraint (2) $v_0\bar{\tau} \geq r_1\tau_0 + \frac{r_2}{2}\bar{\tau}^2$ is binding. If it is not binding, then we increase $r_1$ by $\delta$ and decrease $r_2$ by $\frac{\delta s\theta_0}{\bar{\tau}}$. Constraint (3) is still satisfied. In addition, if $\delta$ is small, (2) is also still satisfied. The objective function changes by

$$
\delta\tau_0 - \frac{\delta s\theta_0}{\bar{\tau}} \frac{\bar{\tau}^2}{2} = \delta\tau_0 \left( 1 - \frac{s\theta_0}{2\bar{\tau}} \right) > 0.
$$

So (2) is binding in an optimal solution. Hence,

$$
v_0\bar{\tau} = r_1\tau_0 + \frac{r_2}{2}\bar{\tau}^2.
$$
Substituting, the profit in the two-part tariff is given by:

$$\Pi_{tpt} = v_0\tilde{\tau} - \frac{r_2}{2}\tilde{\tau}^2 - c\tau_0 + \frac{(r_2 - c)}{2}\tau^2.$$  \hspace{1cm} (18)

If \(\tilde{\tau} < 1\), by (3) and (4), we obtain \(\tilde{\tau} = \frac{v_0}{r_2} - \frac{r_1}{r_2}\theta\tau_0\) and \(\tau = \frac{v_0}{r_2} - \frac{r_1}{r_2}\theta\tau_0\), respectively. Combining them with (17) and after simplification, we get:

\[
\begin{align*}
    r_1 &= \frac{\tilde{\tau}v_0}{2\tau_0 - \theta s\tau_0}; \\
    r_2 &= v_0 \left( \frac{1}{\tau} - \frac{\theta s}{2 - \theta s\tilde{\tau}} \right); \\
    \tau &= \frac{(2 - \theta(s + 1)\tilde{\tau})}{2(1 - \theta s\tau)} \leq \tilde{\tau}.
\end{align*}
\]

It is then straightforward to show that:

\[
\frac{\partial \tau}{\partial \tilde{\tau}} = 2 - \theta(s + 1)\tilde{\tau}(2 - \theta s\tilde{\tau}) \leq 1.
\]

The derivative \(d\Pi_{tpt}/d\tilde{\tau}\) is given by:

\[
\frac{d\Pi_{tpt}}{d\tilde{\tau}} = v_0 - r_2\tilde{\tau} + (r_2 - c)\frac{\partial \tau}{\partial \tilde{\tau}}\tau.
\]  \hspace{1cm} (19)

Substituting for \(r_2\),

\[
v_0 - r_2\tilde{\tau} = \frac{v_0\theta s\tilde{\tau}}{2(1 - \theta s\tilde{\tau}/2)} \geq 0.
\]

It is easy to see that \(v_0 - r_2\tilde{\tau} = \frac{v_0\theta s\tilde{\tau}}{2(1 - \theta s\tilde{\tau}/2)}\) is increasing in \(\tilde{\tau}\). Moreover, \(\frac{\tau}{\tilde{\tau}} \geq 0\), \(\frac{\partial \tau}{\partial \tilde{\tau}} \geq 0\) and it is easy to show that \(r_2\) is decreasing in \(\tilde{\tau}\) and \((r_2 - c)\) becomes negative at some value of \(\tilde{\tau}\). Hence, in (19), there is at most one value of \(\tilde{\tau}^*\) at which \(\frac{d\Pi_{tpt}}{d\tilde{\tau}} = 0\), \(\frac{d\Pi_{tpt}}{d\tilde{\tau}} \geq 0\) for \(\tilde{\tau} \leq \tilde{\tau}^*\), and \(\frac{d\Pi_{tpt}}{d\tilde{\tau}} \leq 0\) for \(\tilde{\tau} > \tilde{\tau}^*\). It means that the profit function \(\Pi_{tpt}\) is unimodal in \(\tilde{\tau}\) and the first order condition is necessary and sufficient for optimality.

Substituting for \(r_2\) in the expression \((r_2 - c)\), we get:

\[
\frac{d\Pi_{tpt}}{d\tilde{\tau}} = \frac{v_0\theta s\tilde{\tau}}{2(1 - \theta s\tilde{\tau}/2)} + v_0\frac{\tau}{\tilde{\tau}}\frac{\partial \tau}{\partial \tilde{\tau}} \left( \frac{1 - \theta s\tilde{\tau} - c\tilde{\tau}/v_0(1 - \theta s\tilde{\tau}/2)}{1 - \theta s\tilde{\tau}/2} \right).
\]  \hspace{1cm} (20)

Hence, the optimal \(\tilde{\tau}\) is obtained by solving the equation:

\[
\frac{v_0\theta s\tilde{\tau}}{2} + v_0\frac{\tau}{\tilde{\tau}}\frac{\partial \tau}{\partial \tilde{\tau}} \left( 1 - \theta s\tilde{\tau} - \frac{c\tilde{\tau}}{v_0(1 - \theta s\tilde{\tau}/2)} \right) = 0.
\]

Hence, the profit function has a unique maximum, which is achieved at \(\tilde{\tau}^* = \text{Min}(\hat{\tau}, 1)\) where \(\hat{\tau}\) is the solution to:

\[
\frac{v_0\theta s\tilde{\tau}}{2} + v_0\frac{\tau}{\tilde{\tau}}\frac{\partial \tau}{\partial \tilde{\tau}} \left( 1 - \theta s\tilde{\tau} - \frac{c\tilde{\tau}}{v_0(1 - \theta s\tilde{\tau}/2)} \right) = 0.
\]
where \( \bar{\tau} = \frac{(1-\theta(\bar{s}+1))\cdot 2}{1-\theta s \bar{\tau}} \) and \( \frac{\partial \bar{\tau}}{\partial \theta} = \frac{2-\theta(\bar{s}+1)\cdot 2 - \theta s \bar{\tau}}{2(1-\theta s \bar{\tau})^2} \).

(ii) If \( \frac{v_0}{c} \geq \frac{1-\theta s \bar{\tau}}{c} \), then \((1-\theta s \bar{\tau}) - \frac{c}{v_0}(1-\theta s \bar{\tau}/2) \geq 0 \) for \( \bar{\tau} \in [0,1] \) and since \( \frac{v_0 \theta s \bar{\tau}}{2} \geq 0 \), \( \frac{d\Pi_{\text{ttl}}(\theta)}{d\bar{\tau}} \geq 0 \) for all \( \bar{\tau} \in [0,1] \) and so \( \bar{\tau}^* = 1 \).

(iii) We now show that the profit increases with \( \theta \). The service provider’s profit \( \Pi_{\text{ttl}}^s(\theta) \) can be written as

\[
\beta^2(v_0\bar{\tau} - \frac{c}{2} \bar{\tau}^2) + (1 - \beta^2\frac{v_0 \bar{\tau}}{\tau_0(2-s\theta \bar{\tau})}),
\]

where \( \beta = \frac{\bar{\tau}}{\bar{\tau}} = \frac{\theta(\bar{s}+1)\cdot 2-\theta s \bar{\tau}}{2(\theta s \bar{\tau} - 2)} \), and \( \beta \) decreases with \( \theta \) and \( \bar{\tau} \) because \( \frac{\partial \beta}{\partial \theta} = \frac{(\bar{s}-1)\cdot 2}{2(\theta s \bar{\tau} - 1)^2} \leq 0 \) and \( \frac{\partial \beta}{\partial \bar{\tau}} = \frac{\theta(\bar{s}-1)}{2(\theta s \bar{\tau} - 1)^2} \leq 0 \).

We prove the result in three steps. In step 1, we show that \( v_0\bar{\tau} - \frac{c}{2} \bar{\tau}^2 \leq \frac{v_0 \bar{\tau}}{\tau_0(2-s\theta \bar{\tau})} \) for \( \bar{\tau} \geq \frac{v_0}{c} \). It is equivalent to show that

\[
v_0 - \frac{c}{2} \bar{\tau} \leq \frac{v_0}{\tau_0(2-s\theta \bar{\tau})}.
\]

We observe that the left hand side of (21) is decreasing in \( \bar{\tau} \), and the right hand side of (21) is increasing in \( \bar{\tau} \). For \( \bar{\tau} = \frac{v_0}{c} \), the above inequality is equivalent to \( \frac{v_0}{2} \leq \frac{v_0}{\tau_0(2-s\theta \bar{\tau})} \). Since \( 2 \geq \tau_0(2-s\theta \bar{\tau}^2), v_0\bar{\tau} - \frac{c}{2} \bar{\tau}^2 \leq \frac{v_0 \bar{\tau}}{\tau_0(2-s\theta \bar{\tau})} \) holds for any \( \bar{\tau} \geq \frac{v_0}{c} \).

In step 2, we show that \( \bar{\tau}^* \geq \min(\frac{v_0}{c}, 1) \). Substituting \( \bar{\tau} = \frac{v_0}{c} \) into the first order condition (20), we have

\[
\frac{d\Pi_{\text{ttl}}(\theta)}{d\bar{\tau}} \bigg|_{\bar{\tau} = \frac{v_0}{c}} = \frac{v_0 \theta s \bar{\tau}}{2} + \frac{v_0 \bar{\tau}}{\tau_0((1-\theta s \bar{\tau}) - (1-\theta s \bar{\tau}/2))} - \frac{v_0 \bar{\tau}}{\tau_0(2-s\theta \bar{\tau})}.
\]

Since \( \bar{\tau} \leq 1 \) and \( \frac{\partial \bar{\tau}}{\partial \theta} \leq 1 \), we get \( \frac{d\Pi_{\text{ttl}}(\theta)}{d\bar{\tau}} \bigg|_{\bar{\tau} = \frac{v_0}{c}} \geq 0 \). Hence, the optimal \( \bar{\tau}^* \) should satisfy that \( \bar{\tau}^* \geq \min(\frac{v_0}{c}, 1) \).

Finally, in step 3, taking the derivative of the objective function \( \Pi_{\text{ttl}}^s(\theta) \) with respect to \( \theta \), we get

\[
\frac{d\Pi_{\text{ttl}}(\theta)}{d\theta} = \frac{\partial \Pi_{\text{ttl}}}{\partial \theta} \bigg|_{\bar{\tau} = \bar{\tau}^*} = 2\beta \frac{\theta}{\theta^*} (v_0\bar{\tau}^* - \frac{c}{2} \bar{\tau}^* \frac{c}{2} - \frac{v_0 \bar{\tau}^*}{\tau_0(2-s\theta \bar{\tau}^*)}) + (1 - \beta^2) \frac{\partial}{\partial \theta} \left( \frac{v_0 \bar{\tau}^*}{\tau_0(2-s\theta \bar{\tau}^*)} \right) \geq 0.
\]

We next show that the service provider’s profit decreases in \( s \). Recall \( \beta = \frac{\theta(\bar{s}+1)\cdot 2-\theta s \bar{\tau}}{2(\theta s \bar{\tau} - 1)^2} \). We have \( \frac{\partial \beta}{\partial \bar{\tau}} = \frac{\theta \bar{s}(1-\theta s \bar{\tau})}{2(\theta s \bar{\tau} - 1)^2} > 0 \). Using envelope theorem, we have that the sensitivity of the service provider’s optimal profit with respect to \( s \) is:

\[
\frac{d\Pi_{\text{ttl}}^s(\theta)}{ds} = \frac{\partial \Pi_{\text{ttl}}^s(\theta)}{\partial \bar{\tau}} \bigg|_{\bar{\tau} = \bar{\tau}^*} = 2\beta \frac{\partial \beta}{\partial \bar{\tau}} (v_0\bar{\tau}^* - \frac{c}{2} \bar{\tau}^* \frac{c}{2} - \frac{v_0 \bar{\tau}^*}{\tau_0(2-s\theta \bar{\tau}^*)}) + (1 - \beta^2) \frac{\partial}{\partial \bar{\tau}} \left( \frac{v_0 \bar{\tau}^*}{\tau_0(2-s\theta \bar{\tau}^*)} \right).
\]

34
Plugging the expressions for $\beta$ and $\frac{\partial \beta}{\partial s}$ into the above equation, we get

$$\frac{\partial \Pi_{tpt}^s(\theta)}{\partial s} = \frac{\theta \tau^s}{4(1 - \theta s \bar{\tau}^s)^2} \left\{ 2[2 - \theta(s+1)\bar{\tau}^s] - \frac{1}{1 - \theta s \bar{\tau}^s} [v_0 \bar{\tau}^s - \frac{c}{2} (\bar{\tau}^s)^2 - \frac{v_0 \bar{\tau}^s}{\tau_0(2 - s \theta \bar{\tau}^s)}] - (4 - 3\theta s \bar{\tau}^s - \theta \bar{\tau}^s) - \frac{v_0 \theta (\bar{\tau}^s)^2}{\tau_0(2 - s \theta \bar{\tau}^s)^2} \right\}.$$

Since it was shown that $v_0 \bar{\tau}^s - \frac{c}{2} (\bar{\tau}^s)^2 - \frac{v_0 \bar{\tau}^s}{\tau_0(2 - s \theta \bar{\tau}^s)} < 0$, we then have $\frac{\partial \Pi_{tpt}^s(\theta)}{\partial s} < 0$.

(iv) We first show that $\frac{\partial r}{\partial \theta} \geq 0$. Since $\frac{\partial \beta}{\partial \theta} \leq 0$ and $\frac{\partial \beta}{\partial s} \leq 0$, we get $\frac{\partial r}{\partial \theta} = \frac{\partial \beta}{\partial \theta} \frac{\partial s}{\partial \theta} \geq 0$. That is, $\bar{\tau}$ increases in $\theta$.

Next, we show that $r_1$ increases in $\theta$. Recall that $r_1 = \frac{\tau v_0}{2 \tau_0 - \theta s \tau}$, when $\theta$ increases, $\bar{\tau}$ increases accordingly. Hence, the numerator increases and the denominator decreases. Therefore, $r_1$ increases in $\theta$.

Finally, we show that $r_2$ decreases in $\theta$. Recall that $r_2 = \frac{v_0}{2} \left( 1 - \frac{s}{2} \bar{\tau}^s \right)$. When $\theta$ increases, $\bar{\tau}$ increases and $\frac{1}{\theta}$ decreases accordingly. Hence, $\frac{\theta s}{\tau^s \bar{\tau}^s}$ increases in $\theta$. So $r_2$ decreases in $\theta$.

The proofs to show how $\bar{\tau}$, $r_1$ and $r_2$ change with $s$ are analogous to those shown above for change with $\theta$, and are therefore skipped.

(b) When customers are sophisticated ($s = 1$), the constraints (3) and (4) are identical and so $\bar{\tau} = \bar{\tau}$. We can first show that $\frac{v_0}{r_2} - \frac{r_1}{r_2} \tau_0 \theta \leq 1$. If not, we can increase $r_2$, and let $r_1 = \frac{v_0}{\tau_0} - \frac{r_1}{r_2}$, the objective function $(r_1 - c) \tau_0 + (r_2 - c) \frac{1}{2}$ increases with respect to $r_2$. Second, solving $\bar{\tau} = \frac{v_0}{r_2} - \frac{r_1}{r_2} \tau_0 \theta$ and $v_0 \bar{\tau} = r_1 \tau_0 + \frac{r_2}{2} \tau^2$, we get $r_1 = \frac{\bar{\tau} v_0}{2 \bar{\tau}_0 - \theta \tau_0 \bar{\tau}}$ and $r_2 = \frac{v_0 2(1 - \theta)}{2 - \theta}$.

Substituting $r_1$ and $r_2$ into the objective function, we have $\Pi_{tpt}^s(\theta) |_{s=1} = -\frac{1}{2} c \left( 2 \tau_0 + \bar{\tau}^2 \right) - v_0 \left( \tau^2 (1 - \theta) s + \bar{\tau}^2 \right)$. The second derivative of the profit function with respect to $\bar{\tau}$ can be simplified to $\frac{d^2 \Pi_{tpt}^s(\bar{\tau})}{d \bar{\tau}^2} = -c$. So the function is concave in $\bar{\tau}$. In addition, $\frac{d \Pi_{tpt}^s(\bar{\tau})}{d \bar{\tau}} |_{s=1} \geq 0$, so $\bar{\tau}^* = 1$, $r_1^* = \frac{v_0}{2 \tau_0 - \theta \tau_0}$ and $r_2^* = \frac{v_0}{2 \tau_0 - \theta \tau_0} \left( \frac{\theta}{\theta - 2} + 1 \right)$; if $v_0 < c$, $\bar{\tau}^* = \frac{v_0}{c}$, $r_1 = \frac{v_0^2 c}{2 \tau_0 c - \theta \tau_0 v_0}$, and $r_2 = v_0 \left( \frac{c \theta}{\tau_0 v_0 - 2 c} + \frac{c}{v_0} \right)$.

**Proof of Proposition 3.** From the proof of Proposition 2, constraint (6) will be binding. So we have $v_0 \bar{\tau} = r_1 \tau_0 (1 - \bar{\tau}) + \frac{r_2}{2} \bar{\tau}^2$. Let $\bar{\tau} = \bar{\tau} = v_0/r_2 < 1$ for the time being. Then, substituting $\bar{\tau} = \bar{\tau} = v_0/r_2$ into $\Pi_{tpt} = r_1 \tau_0 (1 - \bar{\tau}) + (r_2 - c) \frac{\bar{\tau}^2}{2} - c \tau_0$, we get

$\Pi_{tpt} = \frac{v_0^2}{2} \left( \frac{2}{r_2} - \frac{c}{r_2} \right) - c \tau_0$.

Let $x = 1/r_2$. Then, $\Pi_{tpt} = 2x - cx^2$, where $0 \leq x \leq \frac{1}{v_0}$. Is easy to see that the function is concave. From the first order conditions, the optimal solution $x^* = \min(1/c, 1/v_0)$. Hence, $r_2^* = c$ if $v_0/c < 1$. Then, $\bar{\tau}^* = \bar{\tau}^* = v_0/r_2 = v_0/c < 1$. Otherwise, $r_2^* = v_0$, if $v_0/c \geq 1$. Then,
\( \tau^* = \tau^* = v_0/r_2 = 1 \). By substituting \( \tau^* \) and \( \tau^* \) into \( v_0 \tau = r_1\tau_0(1 - \tau) + \frac{\tau^2}{2} \), we can get the optimal value of \( r_1\tau_0 \) when \( v_0/c < 1 \). When \( v_0/c \geq 1 \), customers do not leave after diagnosis. Hence, the fixed fee is not charged to any customer and is then not relevant. The profits follow from substituting the optimal values of \( \tau^* \), \( \tau^* \) and \( r_2^* \) into the objective function \( \Pi_{ctpt} \) in both the scenario that \( v_0 \geq c \) and that \( v_0 \leq c \).

Proof of Proposition 4. (i) The problem formulation is:

\[
\Pi_s^*(\theta) = \max_r (r - c)(\tau_0 + \frac{1}{2}\tau^2)
\]

\[\text{s.t. } v_0 \tau \geq r(\tau_0 + \frac{1}{2}\tau^2), \]

\[\tau = \min(\frac{v_0}{r} - \tau_0 s\theta, 1), \]

\[\tau_0 = \min(\frac{v_0}{r} - \tau_0 \theta, 1), \]

\[r \geq c. \]

First, we have to ensure that \( \tau \in [0, 1] \). That is \( \tau = (\frac{v_0}{r} - \theta s\tau_0) \in [0, 1] \), which gives:

\[r \in \left[\frac{v_0}{1 + \theta s\tau_0}, \frac{v_0}{\theta s\tau_0}\right].\]

Now we substitute \( \tau = \frac{v_0}{r} - \theta s\tau_0 \) into the customer’s participation constraint (at time 0):

\[v_0 \tau \geq r(\tau_0 + \frac{1}{2}\tau^2),\]

which gives

\[r \leq \bar{\tau} = \frac{v_0}{\sqrt{k}},\]

where \( k := 2\tau_0 + \theta^2 s^2 \tau_0^2 \). We now have two cases: 1) \( \frac{v_0}{\sqrt{k}} \geq c \); and 2) \( \frac{v_0}{\sqrt{k}} < c \).

- Case 1: \( \frac{v_0}{\sqrt{k}} \geq c \). In this case, the range of \( r \) we are searching for is: \( r \in [c, \frac{v_0}{\sqrt{k}}] \).

- Case 2: When \( \frac{v_0}{\sqrt{k}} < c \), the service provider makes negative profit by choosing \( r = \frac{v_0}{\sqrt{k}} < c \).

Even if the provider chooses \( r = c \), the customers will not join in the service because the constraint \( v_0 \tau \geq r(\tau_0 + \frac{1}{2}\tau^2) \) is violated. So the market is inactive.

Next, we analyze the case such that \( \frac{v_0}{\sqrt{k}} \geq c \) and correspondingly \( r \in [c, \frac{v_0}{\sqrt{k}}] \) henceforth.
The provider’s optimization problem can then be rewritten as:

\[
\max_r (r - c)[\tau_0 + \frac{1}{2}(\frac{v_0}{r} - \theta \tau_0)^2] \\
\text{s.t. } c \leq r \leq \frac{v_0}{\sqrt{k}}.
\]

Consider the objective function

\[\Pi_t^\tau(\theta) := (r - c)[\tau_0 + \frac{1}{2}(\frac{v_0}{r} - \theta \tau_0)^2],\]

where \(r \in [c, \frac{v_0}{\sqrt{k}}]\). The first derivative of \(\Pi_t^\tau\) with respect to \(r\) is

\[
\frac{\partial \Pi_t^\tau(\theta)}{\partial r} = \frac{2cv_0(v_0 - \theta r \tau_0) + 3\tau_0(r^2 \tau_0 + 2) - rv_0^2}{2v_0^2}.
\]

Note that it is easy to show that \(\frac{\partial \Pi_t^\tau(\theta)}{\partial r} \bigg|_{r = \frac{v_0}{\sqrt{k}}} = \tau_0 \frac{2c(v_0^2 \tau_0^2 + 2)(\sqrt{\tau_0^2 (v_0^2 + 2) - \theta \tau_0 - \theta^2 (s^2 - 1) \tau_0 v_0})}{2v_0^2} > 0\)

and \(\frac{\partial \Pi_t^\tau(\theta)}{\partial r} \bigg|_{r = c} = \frac{1}{2} \left( \frac{v_0}{c} - \theta \tau_0 \right)^2 + \tau_0 > 0\). In addition, we can get that

\[
\frac{\partial^2 \Pi_t^\tau(\theta)}{\partial r^2} = \frac{v_0}{r^3} \left( -\frac{3cv_0}{r} + v_0 + 2c\theta \tau_0 \right).
\]

Let \(\hat{r}\) be the value of \(r\) such that \(\frac{\partial^2 \Pi_t^\tau(\theta)}{\partial r^2} = 0\). Note that \(\frac{\partial^2 \Pi_t^\tau(\theta)}{\partial r} \geq 0\) for \(r \geq \hat{r}\); and \(\frac{\partial^2 \Pi_t^\tau(\theta)}{\partial r} < 0\) for \(r < \hat{r}\).

i) If \(\hat{r} \geq \frac{v_0}{\sqrt{k}}\), \(\frac{\partial^2 \Pi_t^\tau(\theta)}{\partial r} \geq 0\) for \(r \in [\hat{r}, \frac{v_0}{\sqrt{k}}]\). \(\Pi_t^\tau\) is a concave function with \(\frac{\partial \Pi_t^\tau(\theta)}{\partial r} \bigg|_{r = \frac{v_0}{\sqrt{k}}} > 0\). Thus \(\Pi_t^\tau(\theta)\) monotone increases in \(r\) since \(\frac{\partial \Pi_t^\tau(\theta)}{\partial r}\) decreases yet remains positive in this range of \(r\). So \(r^* = \frac{v_0}{\sqrt{k}}\).

ii) If \(\hat{r} \leq c\), \(\frac{\partial^2 \Pi_t^\tau(\theta)}{\partial r^2} \geq 0\) for \(r \in [c, \frac{v_0}{\sqrt{k}}]\). \(\Pi_t^\tau\) is a convex function with \(\frac{\partial \Pi_t^\tau(\theta)}{\partial r} \bigg|_{r = \frac{v_0}{\sqrt{k}}} > 0\) and \(\frac{\partial \Pi_t^\tau(\theta)}{\partial r} \bigg|_{r = c} > 0\). So \(r^* = \frac{v_0}{\sqrt{k}}\).

iii) If \(c \leq \hat{r} \leq \frac{v_0}{\sqrt{k}}\) and \(\frac{\partial \Pi_t^\tau(\theta)}{\partial r} \bigg|_{r = \frac{v_0}{\sqrt{k}}} = \frac{v_0}{c^2} + \frac{v_0}{c^2}(v_0 - c\theta \tau_0) \geq 0\), \(\frac{\partial \Pi_t^\tau(\theta)}{\partial r}\) decreases and then increases in the range of \(r \in [c, \frac{v_0}{\sqrt{k}}]\), and the only turning point is exactly at \(\hat{r}\). Recall that we had \(\frac{\partial \Pi_t^\tau(\theta)}{\partial r} \bigg|_{r = \frac{v_0}{\sqrt{k}}} \geq 0\), and \(\frac{\partial \Pi_t^\tau(\theta)}{\partial r} \bigg|_{r = \frac{v_0}{\sqrt{k}}} > 0\). The uniqueness about the turning point \(\hat{r}\) leads to the fact that \(\frac{\partial \Pi_t^\tau(\theta)}{\partial r} > 0\) in this entire range of \(r\). So, \(r^* = \frac{v_0}{\sqrt{k}}\).

To summarize, we have that \(r^* = \frac{v_0}{\sqrt{k}}\). Note that \(\tau^* = \frac{v_0}{r} - \theta s \tau_0 = \sqrt{k} - \theta s \tau_0\). Substituting for \(\tau^*\) and \(r^*\) in the profit function, we obtain the expression for the optimal profit. This concludes the proof.

(ii) From the proof of part (i), we get \(\frac{\partial \Pi_t^\tau(\theta)}{\partial \theta} \leq 0\). Thus, the service provider’s profit decreases in \(\theta\). In addition, the service provider’s problem can be reduced to \((r - c)[\tau_0 + \frac{1}{2}(\frac{v_0}{r} - \theta \tau_0)^2]\) given the constraint that \(r \leq \frac{\frac{v_0}{\sqrt{2\tau_0 + \theta^2 s^2 \tau_0^2}}}{\sqrt{2\tau_0 + \theta^2 s^2 \tau_0^2}}\). The feasible region is larger when \(s\) is small. So the service provider’s profit decreases in \(s\).
**Proof of Proposition 5.** The formulation of the conditional fixed fee problem is:

\[
\Pi_{cf} = \max_{\{f, \bar{\tau}\}} \quad f \bar{\tau} - \frac{\bar{\tau}^2}{2} - c\tau_0
\]

subject to:

\[
v_0 \bar{\tau} \geq f \bar{\tau}
\]

\[
\bar{\tau} \in (0, 1).
\]

Recall that under the unconditional fixed fee scheme, \(\Pi_f = v_0 - \frac{c}{2} - c\tau_0\). So, if \(v_0 \geq c\), there is no additional profit since all customers are served in both schemes and \(f^* = v_0\). So, we consider the case that \(v_0 < c\).

Suppose the provider posts a conditional fixed fee, i.e., customers are uniformly charged \(f\) upon joining the service and the provider has the right to reject customers. It is clear that the provider under this new payment scheme will reject treatment for customers only when their type \(\tau > \bar{\tau}\) upon diagnosis. So, the provider solves

\[
\max_{f, \bar{\tau}} \quad f \bar{\tau} - \frac{\bar{\tau}^2}{2} - c\tau_0
\]

subject to:

\[
v_0 \bar{\tau} \geq f \bar{\tau}
\]

\[
\bar{\tau} \in [0, 1].
\]

The first constraint is the customer’s participation condition, which should be binding at optimality. Consequently, \(f^* = v_0\).

Plugging \(v_0 \bar{\tau} = f \bar{\tau}\) into the objective function gives

\[
\max_{\{\bar{\tau}\}} \quad v_0 \bar{\tau} - \frac{\bar{\tau}^2}{2} - c\tau_0,
\]

which yields \(\bar{\tau}^* = \min(\frac{v_0}{c}, 1)\).

In summary, \((f^*, \bar{\tau}^*) = (v_0, \frac{v_0}{c})\) and the service provider’s profit is \(\Pi_{cf} = \frac{v_0^2}{c} - \frac{v_0^2}{2c} - c\tau_0 = \frac{v_0^2}{2c} - c\tau_0\) when \(v_0 < c\); \((f^*, \bar{\tau}^*) = (v_0, 1)\) and the provider’s profit is \(\Pi_{cf} = v_0 - \frac{1}{2}c - c\tau_0\) when \(v_0 \geq c\). Recall that under an unconditional or simple fixed fee, \(\Pi_f = v_0 - \frac{c}{2} - c\tau_0\) when \(v_0 \geq c\). Thus, if \(v_0 \geq c\), \(\Pi_{cf} = \Pi_f\). If \(v_0 < c\), the additional profit to the provider by rejecting customers with high treatment time is:

\[
\Pi_{cf} - \Pi_f = \frac{v_0^2}{2c} - v_0 + \frac{c}{2} = \frac{(v_0 - c)^2}{2c}.
\]

We find that the profit under this payment scheme is the same as the profit under the two-part tariff when \(\theta = 0\) (in Proposition 1) or \(s = 0\) (in part (b) of Proposition 2).
Proof of Proposition 6: If $r_1 = 0$, the problem becomes

$$\Pi_2 = \max_{r_2} \quad -c\tau_0 + (r_2 - c)\frac{r_2^2}{2}$$

$$s.t. \quad v_0\bar{r} \geq \frac{r_2 + c}{2},$$

$$\bar{r} = \bar{r} = \min(\frac{v_0}{r_2}, 1).$$

(25)

If $v_0 \leq r_2$, $\bar{r} = \bar{r} = \frac{v_0}{r_2}$, and the objective becomes $-c\tau_0 + (r_2 - c)\frac{v_0^2}{2r_2^2}$. The optimal rate $r_2^* = 2c$ and the optimal profit is $\frac{v_0^2}{8c} - c\tau_0$. It is valid when $v_0 \leq 2c$.

If $v_0 \geq r_2$, $\bar{r} = \bar{r} = 1$, then $r_2 \leq 2v_0$ and the objective is $-c\tau_0 + (r_2 - c)\frac{1}{2}$, which increases with $r_2$. So $r_2 = v_0$ and the profit is $\frac{1}{2}v_0 - c\left(\frac{1}{2} + \tau_0\right)$ when $v_0 \geq 2c$. The profit is lower than the profit under optimal fixed fee which equals $v_0 - c\left(\frac{1}{2} + \tau_0\right)$.

Proof of Proposition 7. (a) From Proposition 5, if $v_0 < c$, $\Pi_{cf} = \frac{v_0^2}{2c} - c\tau_0$. Otherwise, if $v_0 \geq c$, $\Pi_{cf} = v_0 - \frac{c}{2} - c\tau_0$. We know that $\Pi_f = v_0 - \frac{c}{2} - c\tau_0$.

If $v_0 < c$, then $\frac{v_0^2}{2c} \geq v_0 - \frac{c}{2}$. Hence $\Pi_{cf} \geq \Pi_f$. If $v_0 \geq c$, then $\Pi_{cf} = \Pi_f$. As a result, $\Pi_{cf} \geq \Pi_f$.

Next, we show that $\Pi_{cf} \geq \Pi_{fd}$. From Proposition 6, when $v_0 \geq 2c$, $\Pi_{fd} = \frac{1}{2}v_0 - c(\tau_0 + \frac{1}{2}) \leq \Pi_f = \Pi_{cf}$, and when $v_0 \leq 2c$, $\Pi_{fd} = \frac{1}{8}v_0^2 - c\tau_0 \leq \Pi_{cf} = \frac{v_0^2}{2c} - c\tau_0$. Hence, $\Pi_{cf} \geq \Pi_{fd}$.

Now we show that $\Pi_{cf} \geq \Pi^*_f(\theta)$.

From Proposition 5, if $v_0 < c$, $\Pi_{cf} = \frac{v_0^2}{2c} - c\tau_0$. Otherwise, if $v_0 \geq c$, $\Pi_{cf} = v_0 - \frac{c}{2} - c\tau_0$.

From Proposition 4, part (i), we know that $\Pi^*_f(\theta) = \left(\frac{v_0}{\sqrt{8c}} - c\right)(\tau_0 + \frac{1}{2}(\sqrt{8} - \theta \tau_0)^2)$. Note that $\Pi^*_f(\theta) \geq 0$ requires $\frac{v_0}{\sqrt{8c}} \geq c$. From Proposition 4, profit decreases with $\theta$. Next, we show that $\Pi_{cf} \geq \Pi^*_f(\theta = 0)$ which implies that $\Pi_{cf} \geq \Pi^*_f(\theta) \forall \theta$. Note that $\Pi^*_f(\theta = 0) = \left(\frac{v_0}{\sqrt{8c}} - c\right)(\tau_0 + \frac{1}{2}(\sqrt{8})^2) = (v_0\sqrt{2\tau_0} - c\tau_0 - c\theta_0)$ after some simplification.

We first consider the case $v_0 < c$.

To show that $\Pi_{cf} \geq \Pi^*_f(\theta = 0)$, we need to show that $\frac{v_0^2}{2c} \geq v_0\sqrt{2\tau_0} - c\tau_0$. After some simplification, this is equivalent to $(v_0 - c\sqrt{2\tau_0})^2 \geq 0$. This is valid when $v_0 \geq c\sqrt{2\tau_0}$, an inequality that is satisfied according to Proposition 4, part (i).

Now consider the case $v_0 \geq c$.

To show that $\Pi_{cf} \geq \Pi^*_f(\theta = 0)$ we need to show that $v_0 - \frac{c}{2} \geq v_0\sqrt{2\tau_0} - c\tau_0$. After some simplification, this is equivalent to requiring that $v_0(1 - \sqrt{2\tau_0}) \geq c(\frac{1}{2} - \tau_0)$. Since $v_0 \geq c$, the inequality will be satisfied if $(1 - \tau_0) \geq (\frac{1}{2} - \tau_0)$. After some rearrangement and simplification, this is equivalent to the requirement $(\tau_0 - \frac{1}{2})^2 \geq 0$ which is always true.

(b) From Proposition 6, when $v_0 \geq 2c$, $\Pi_{fd} = \frac{1}{2}v_0 - c(\tau_0 + \frac{1}{2}) \leq \Pi_f$. When $v_0 \leq 2c$, $\Pi_f = v_0 - c(\tau_0 + \frac{1}{2})$ and $\Pi_{fd} = \frac{1}{8}v_0^2 - c\tau_0$. Let $v_0 = kc$. To show $\Pi_f \geq \Pi_{fd}$, we then require $kc - \frac{c}{2} \geq \frac{k^2c}{8}$ or $\frac{k^2}{8} - k + \frac{1}{2} \geq 0$ which translates to $k \geq 0.536$. Hence, $\Pi_f \geq \Pi_{fd}$ if $v_0 \geq 0.536c$. 39
(c) By the result of Proposition 4, part (i), the service provider’s profit under the single rate time-based fee is: \( \Pi^s(\theta) = \left( \frac{\nu_0}{\sqrt{2t_0 + \theta^2 s^2 t_0^2}} - c \right)(\tau_0 + \frac{1}{2} \tau^2) = \left( \frac{\nu_0}{\sqrt{2t_0 + \theta^2 s^2 t_0^2}} - c \right)(\tau_0 + \frac{1}{2} \left( \sqrt{2t_0} + \theta^2 s^2 t_0^2 - \theta \tau_0 \right)^2). \) Taking the derivative of profit function \( \Pi^s(\theta) \) with respect to \( \theta \), we get,

\[
\frac{\partial \Pi^s(\theta)}{\partial \theta} = \tau_0 \left( \frac{\theta s^2 \tau_0}{\sqrt{\tau_0 (\theta^2 s^2 \tau_0 + 2)}} - 1 \right) \left( \sqrt{\tau_0 (\theta^2 s^2 \tau_0 + 2)} - \theta \tau_0 \right) \left( \frac{\nu_0}{\sqrt{\tau_0 (\theta^2 s^2 \tau_0 + 2)}} - c \right) - \frac{\theta s^2 \tau_0 \nu_0}{\left( \tau_0 (\theta^2 s^2 \tau_0 + 2) \right)^{3/2}} \left( \frac{1}{2} \left( \sqrt{\tau_0 (\theta^2 s^2 \tau_0 + 2)} - \theta \tau_0 \right)^2 + \tau_0 \right) \leq 0.
\]

The inequality is true because \( \theta s^2 \tau_0 \leq \sqrt{\tau_0 (\theta^2 s^2 \tau_0 + 2)} \), \( \sqrt{\tau_0 (\theta^2 s^2 \tau_0 + 2)} - \theta \tau_0 \geq 0 \) and \( \frac{\nu_0}{\sqrt{\tau_0 (\theta^2 s^2 \tau_0 + 2)}} \geq c. \)

When \( \theta = 0 \) and \( \theta = 1 \), we have

\[ \Pi^s(\theta = 0) = 2\tau_0 \left( \frac{\nu_0}{\sqrt{2\tau_0}} - c \right) \]

and

\[ \Pi^s(\theta = 1) = \left( \frac{\nu_0}{\sqrt{s^2 \tau_0^2 + 2\tau_0}} - c \right) \left( \tau_0 + \frac{1}{2} \left( \sqrt{s^2 \tau_0^2 + 2\tau_0 - \tau_0} \right)^2 \right), \]

respectively.

Let \( L(s, \tau_0) = \frac{(\tau_0 - \sqrt{\tau_0 (s^2 \tau_0 + 2)})^2 - 1}{2 \left( \frac{1}{2} (\tau_0 - \sqrt{\tau_0 (s^2 \tau_0 + 2)})^2 + \tau_0 \right)^2}, \) and \( U(s, \tau_0) = \frac{1 + \sqrt{2\tau_0}}{2}. \)

There are three possible scenarios that can happen:

(i) If \( \Pi^s(1) \leq \Pi_f \leq \Pi^s(0) \), that is, \( L(s, \tau_0) \leq \frac{\nu_0}{c} \leq U(s, \tau_0) \), there exists \( \bar{s} \in [0, 1] \) such that \( \Pi^s(\theta) \geq \Pi_f \) if \( \theta \leq \bar{s} \) and \( \Pi^s(\theta) < \Pi_f \) if \( \theta > \bar{s} \). \( \bar{s} \) solves \( \Pi^s(\theta) = \Pi_f \). Since \( \frac{\partial \Pi^s(\theta)}{\partial \theta} = \frac{\theta^2 s^3 (s^2 - 1) \tau_0 \nu_0 - 2c(\theta^2 s^2 \tau_0 + 2) \left( \sqrt{\tau_0 (\theta^2 s^2 \tau_0 + 2)} - \theta \tau_0 \right)}{2(\tau_0 (\theta^2 s^2 \tau_0 + 2))^{3/2}} \leq 0 \) for \( s \in [0, 1] \), we get \( \bar{s} \) decreases with the sophistication level \( s \).

(ii) If \( \Pi_f \geq \Pi^s(0) \), i.e. \( \frac{\nu_0}{c} \geq U(s, \tau_0) \), then \( \Pi_f \geq \Pi^s(\theta) \) for \( \theta \in [0, 1] \), and fixed scheme is the dominant strategy.

(iii) Otherwise, if \( \Pi_f \leq \Pi^s(1) \), i.e. \( L(s, \tau_0) > \frac{\nu_0}{c} \), then \( \Pi_f \leq \Pi^s(\theta) \) for \( \theta \in [0, 1] \), and time-based scheme is the dominant strategy. This concludes the proof. \( \blacksquare \)

**Proof of Proposition 8:** We first show that the service provider cannot be better off by offering two different two part tariff to customers with heterogeneous sunk cost bias \( \theta_L \) and \( \theta_H \). Assume that the service provider offers two different two part tariff: \( (r_{1H}, r_{2H}) \) and \( (r_{1L}, r_{2L}) \) to attract \( \theta_H \) type and \( \theta_L \) type, respectively. Then, the contracts \( (r_{1H}, r_{2H}) \) and \( (r_{1L}, r_{2L}) \) should solve the
The following problems:

\[
\begin{align*}
\max_{r_{1H}, r_{2H}, r_{1L}, r_{2L}} & \quad \lambda((r_{1H} - c)\tau_0 + (r_{2H} - c)\frac{1}{2}\tau^2_H) + (1 - \lambda)((r_{1L} - c)\tau_0 + (r_{2L} - c)\frac{1}{2}\tau^2_L) \\
\text{s.t.} & \quad v_0 \tau_H \geq r_{1H}\tau_0 + \frac{r_{2H}^2}{2}\tau^2_H \\
& \quad v_0 \tau_L \geq r_{1L}\tau_0 + \frac{r_{2L}^2}{2}\tau^2_L \\
& \quad v_0 \tau_H - (r_{1H}\tau_0 + \frac{r_{2H}^2}{2}\tau^2_H) \geq v_0 \tau^L_H - (r_{1L}\tau_0 + \frac{r_{2L}^2}{2}(\tau^L_H)^2) \\
& \quad v_0 \tau_L - (r_{1L}\tau_0 + \frac{r_{2L}^2}{2}\tau^2_L) \geq v_0 \tau^L_H - (r_{1H}\tau_0 + \frac{r_{2H}^2}{2}(\tau^H_L)^2) \\
& \quad \tau^L_H = \min\left(\frac{v_0}{r_{2L}} - \frac{r_{1L}}{r_{2L}}\theta^L_H, 1\right) \\
& \quad \tau^L_L = \min\left(\frac{v_0}{r_{2L}} - \frac{r_{1L}}{r_{2L}}\theta^L_L, 1\right) \\
& \quad \tau^H_H = \min\left(\frac{v_0}{r_{2H}} - \frac{r_{1H}}{r_{2H}}\theta^H_H, 1\right) \\
& \quad \tau^H_L = \min\left(\frac{v_0}{r_{2H}} - \frac{r_{1L}}{r_{2H}}\theta^H_L, 1\right)
\end{align*}
\]

The constraints (28) and (29) are the participation constraints for \(\theta_H\) type and \(\theta_L\) type customers before starting the diagnosis, respectively. The constraint (30) makes sure that \(\theta_H\) type customers do not want to choose \((r_{1L}, r_{2L})\), and the constraint (31) makes sure that \(\theta_L\) type customers do not want to choose \((r_{1H}, r_{2H})\). \(\tau^L_H\) in the equation (36) represents the perceived proportion of \(\theta_H\) type customers who continue with the service after diagnosis when choosing the contract \((r_{1L}, r_{2L})\). Similarly, \(\tau^H_L\) in the equation (37) represents the perceived proportion of \(\theta_L\) type customers who continue with the service after diagnosis when choosing the contract \((r_{1H}, r_{2H})\). We first ignore the fact that \(\tau_H\) and \(\tau_L\) could be equal to 1. Then, the optimization problem (27) – (37) can be
simplified to the following:

\[
\max_{r_{1H}, r_{2H}, r_{1L}, r_{2L}} \lambda ((r_{1H} - c)\tau_0 + (r_{2H} - c)\frac{1}{2}r_{1H}^2) + (1 - \lambda)((r_{1L} - c)\tau_0 + (r_{2L} - c)\frac{1}{2}r_{1L}^2)
\]  

(38)

\[
s.t. \quad v_0^2 \geq 2r_{1H}r_{2H}\tau_0 + r_{1H}^2\theta_H^2s^2\tau_0^2
\]  

(39)

\[
v_0^2 \geq 2r_{1L}r_{2L}\tau_0 + r_{1L}^2\theta_L^2s^2\tau_0^2
\]  

(40)

\[
\frac{v_0^2}{r_{2H}} - 2r_{1H}\tau_0 - \theta_H^2s^2\tau_0^2r_{1H}^2 \geq \frac{v_0^2}{r_{2L}} - 2r_{1L}\tau_0 - \theta_H^2s^2\tau_0^2r_{1L}^2
\]  

(41)

\[
\frac{v_0^2}{r_{2L}} - 2r_{1L}\tau_0 - \theta_L^2s^2\tau_0^2r_{1L}^2 \geq \frac{v_0^2}{r_{2H}} - 2r_{1H}\tau_0 - \theta_L^2s^2\tau_0^2r_{1H}^2
\]  

(42)

\[
\frac{v_0}{r_{2L}} - \frac{r_{1L}}{r_{2L}}\theta_Ls\tau_0 \leq 1
\]  

(43)

\[
\sum_L = \min(\frac{v_0}{r_{2L}} - \frac{r_{1L}}{r_{2L}}\theta_Ls\tau_0, 1)
\]  

(44)

\[
\frac{v_0}{r_{1L}} - \frac{r_{1H}}{r_{2L}}\theta_Hs\tau_0 \leq 1
\]  

(45)

\[
\sum_H = \min(\frac{v_0}{r_{1L}} - \frac{r_{1H}}{r_{2L}}\theta_Hs\tau_0, 1)
\]  

(46)

Suppose either \(r_{1L} \leq r_{1H}\) and \(r_{2L} \geq r_{2H}\), or \(r_{1L} \geq r_{1H}\) and \(r_{2L} \leq r_{2H}\); one could verify that (30) and (31) cannot be true at the same time unless \(r_{1L} = r_{1H}\) and \(r_{2L} = r_{2H}\). In addition, if either \(r_{1L} \leq r_{1H}\) and \(r_{2L} \leq r_{2H}\), or \(r_{1L} \geq r_{1H}\) and \(r_{2L} \geq r_{2H}\), both types of customers will prefer the same contract. As a result, the service provider will achieve the same profit by offering one contract. In the following, we consider the case that the service provider offers one contract \((r_1, r_2)\).

The problem that the service provider needs to solve is:

\[
\max_{r_1, r_2} \lambda((r_1 - c)\tau_0 + (r_2 - c)\frac{1}{2}\tau_H^2) + (1 - \lambda)((r_1 - c)\tau_0 + (r_2 - c)\frac{1}{2}\tau_L^2)
\]  

(47)

\[
s.t. \quad v_0^2 \geq 2r_1r_2\tau_0 + r_1^2\theta_H^2s^2\tau_0^2
\]  

(48)

\[
v_0^2 \geq 2r_1r_2\tau_0 + r_1^2\theta_L^2s^2\tau_0^2
\]  

(49)

\[
\sum_L = \min(\frac{v_0}{r_2} - \frac{r_1}{r_2}\theta_Ls\tau_0, 1)
\]  

(50)

Either (47) or (48) will be binding at optimality. Let \((r_1^*, r_2^*H)\) and \((r_1^*, r_2^*L)\) be the optimal two part tariff to offer when serving only \(\theta_H\) type and only \(\theta_L\) type customers, respectively. Since \(\theta_H > \theta_L\), the service provider either serves only low type customers by offering \((r_1^*, r_2^*L)\), or serves both types of customers by offering \((r_1^*, r_2^*H)\). Following a similar logic, we could show that the results hold when \(\tau_H = 1\) or \(\tau_L = 1\). This concludes the proof. \(\blacksquare\)
Proof of Proposition 9  Let \((r_{1\theta}^*, r_{2\theta}^*)\) be the optimal solution when the provider treats the two types as a single type with \(\theta = \lambda \theta_H + (1 - \lambda) \theta_L\). \((r_{1\theta}^*, r_{2\theta}^*)\) satisfies the following equations:

\[
v_0 \bar{\tau} = r_1 \tau_0 + \frac{r_2^2}{2}\bar{\tau}^2
\]

(51)

\[
\bar{\tau} = \min\left(\frac{v_0}{r_2} - \frac{r_1}{r_2} \theta s \tau_0, 1\right)
\]

(52)

\[
\bar{\tau} = \min\left(\frac{v_0}{r_2} - \frac{r_1}{r_2} \theta_0, 1\right)
\]

(53)

(51) can be simplified to \(v_0^2 \geq 2r_1 r_2 \tau_0 + \frac{r_1^2}{r_2^2} \theta^2 s^2 \tau_0^2\) for \(\frac{v_0}{r_2} - \frac{r_1}{r_2} \theta s \tau_0 \leq 1\). Given that \((r_{1H}^*, r_{2H}^*)\) is the optimal two part tariff to offer when serving only \(\theta_H\) type, we have

\[
v_0^2 = 2r_{1H}^* r_{2H}^* \tau_0 + r_{1H}^* \theta_{H}^2 s^2 \tau_0^2 \geq 2r_{1H}^* r_{2H}^* \tau_0 + r_{1H}^* \theta_{H}^2 s^2 \tau_0^2.
\]

Given that \((r_{1L}^*, r_{2L}^*)\) is the optimal two part tariff to offer when serving only \(\theta_L\) type, we have

\[
v_0^2 = 2r_{1L}^* r_{2L}^* \tau_0 + r_{1L}^* \theta_{L}^2 s^2 \tau_0^2 \leq 2r_{1L}^* r_{2L}^* \tau_0 + r_{1L}^* \theta_{L}^2 s^2 \tau_0^2.
\]

In addition, if \((r_{1\theta}^*, r_{2\theta}^*)\) is the optimal solution when treating customers as homogeneous type \(\theta\), the constraint (51) is binding. Thus,

\[
v_0^2 = 2r_{1\theta}^* r_{2\theta}^* \tau_0 + r_{1\theta}^* \theta^2 s^2 \tau_0^2.
\]

Since \(\theta_L \leq \theta \leq \theta_H\), it implies that

\[
2r_{1\theta}^* r_{2\theta}^* \tau_0 + r_{1\theta}^* \theta_{H}^2 s^2 \tau_0^2 \geq 2r_{1\theta}^* r_{2\theta}^* \tau_0 + r_{1\theta}^* \theta_{H}^2 s^2 \tau_0^2 = v_0
\]

\[
2r_{1\theta}^* r_{2\theta}^* \tau_0 + r_{1\theta}^* \theta_{L}^2 s^2 \tau_0^2 \leq 2r_{1\theta}^* r_{2\theta}^* \tau_0 + r_{1\theta}^* \theta_{L}^2 s^2 \tau_0^2 = v_0.
\]

That is, \(v_0 \bar{\tau} < r_{1\theta}^* \tau_0 + \frac{r_{1\theta}^2}{2} \theta_{H}^2 r_{2\theta}^2\) and \(v_0 \bar{\tau} > r_{1\theta}^* \tau_0 + \frac{r_{1\theta}^2}{2} \theta_{L}^2 r_{2\theta}^2\). Hence, the \(\theta_H\) type customer will not join the service. The firm serves only \(\theta_L\) type customers and leaves positive surplus to them. The total profit in this case is

\[
(1 - \lambda)((r_{1\theta}^* - c) \tau_0 + (r_{2\theta}^* - c) \frac{1}{2} \theta_{L}^2 r_{2\theta}^2) = (1 - \lambda)((r_{1\theta}^* - c) \tau_0 + (r_{2\theta}^* - c) \frac{1}{2} \theta_{L}^2 r_{2\theta}^2).
\]

This concludes the proof. ■

43
Proof of Proposition 10: If the service provider serves both types of customers, the problem he is faced with is:

\[
\max_{r_1, r_2} (r_1 - c) \tau_0 + (r_2 - c) \frac{1}{2} \tau_2^2 \\
\text{s.t. } v_0 \tau_H \geq r_1 \tau_0 + \frac{r_2}{2} \tau_2^2 \\
v_0 \tau_L \geq r_1 \tau_0 + \frac{r_2}{2} \tau_2^2 \\
\tau_L = \min\left(\frac{v_0}{r_2} - \frac{r_1}{r_2} \theta \tau_0, 1\right) \\
\tau_H = \min\left(\frac{v_0}{r_1} - \frac{r_1}{r_2} \theta \tau_0, 1\right) \\
\tau = \min\left(\frac{v_0}{r_2} - \frac{r_1}{r_2} \theta \tau_0, 1\right). 
\]

On the other hand, if the service provider serves only \(s_L\) type customers, the formulation is:

\[
\max_{r_1, r_2} (1 - \gamma)((r_1 - c) \tau_0 + (r_2 - c) \frac{1}{2} \tau_2^2) \\
\text{s.t. } v_0 \tau_L \geq r_1 \tau_0 + \frac{r_2}{2} \tau_2^2 \\
\tau_L = \min\left(\frac{v_0}{r_2} - \frac{r_1}{r_2} \theta \tau_0, 1\right) \\
\tau = \min\left(\frac{v_0}{r_2} - \frac{r_1}{r_2} \theta \tau_0, 1\right). 
\]

The proof is similar to the proof of Proposition 8 and is therefore omitted.

Proof of Proposition 11: The proof is similar to the proof of Proposition 9 and is therefore omitted.
Appendix B: Model with value increasing with treatment time

We consider the analysis of two part tariff where value increases linearly with time. Let \( v(\tau) = v_0 + \kappa \tau \), where \( \kappa > 0 \).

1) At time 0, a customer believes that he will join and complete the service process if and only if \( \tau < \bar{\tau} \), where,

\[
\begin{align*}
v_0 + \kappa \bar{\tau} - r_1 \theta s \tau_0 &= r_2 \bar{\tau}, \text{i.e.} v_0 - r_1 \theta s \tau_0 = (r_2 - \kappa) \bar{\tau}.
\end{align*}
\]

So, a customer’s participation condition can be expressed as:

\[
\int_0^{\bar{\tau}} (v_0 + \kappa \tau) d\tau \geq r(\tau_0 + \frac{\bar{\tau}}{2}) \Rightarrow v_0 \bar{\tau} \geq r_1 \tau_0 + \frac{(r_2 - \kappa) \bar{\tau}^2}{2}.
\]

2) At time \( \tau_0 \), the customer will continue with the service process after diagnosis if and only if \( \tau < \bar{\tau} \), where

\[
\begin{align*}
v_0 + \kappa \bar{\tau} - r_1 \theta \tau_0 &= r_2 \bar{\tau} \Rightarrow v_0 - r_1 \theta \tau_0 = (r_2 - \kappa) \bar{\tau}.
\end{align*}
\]

Hence, the provider solves the problem:

\[
\max_{\hat{r}} \quad \Pi_t = (r_1 - c)\tau_0 + (r_2 - c)\bar{\tau}^2 \quad \text{s.t.} \quad v_0 \bar{\tau} \geq r_1 \tau_0 + \frac{(r_2 - \kappa) \bar{\tau}^2}{2}
\]

\[
\bar{\tau} = \min(\frac{v_0 - r_1 \theta s \tau_0}{r_2 - \kappa}, 1)
\]

\[
\tau = \min(\frac{v_0 - r_1 \theta \tau_0}{r_2 - \kappa}, 1)
\]

\[
\bar{\tau} \geq 0, \tau \geq 0.
\]  

(62)

Recall the provider’s optimization problem under the time-based scheme when we assume a constant value \( v_0 \) (effectively, \( \kappa = 0 \)) is:

\[
\max_{\hat{r}} \quad \Pi_t = (r_1 - c)\tau_0 + (r_2 - c)\bar{\tau}^2 \quad \text{s.t.} \quad v_0 \bar{\tau} \geq r_1 \tau_0 + \frac{(r_2 - \kappa) \bar{\tau}^2}{2}
\]

\[
\bar{\tau} = \min(\frac{v_0 - r_1 \theta s \tau_0}{r_2 - \kappa}, 1)
\]

\[
\tau = \min(\frac{v_0 - r_1 \theta \tau_0}{r_2 - \kappa}, 1)
\]

\[
\bar{\tau} \geq 0, \tau \geq 0.
\]  

(65)

Let \( \hat{r}_2 = r_2 - \kappa, \hat{c} = c - \kappa \). Then, the provider’s problem when \( v(\tau) = v_0 + \kappa \tau \) can be reduced
\[
\max_r \quad \Pi_t = (r_1 - \hat{c})\tau_0 + (\hat{r}_2 - \hat{c})\frac{\tau^2}{2} - \kappa \tau_0
\]

s.t \quad v_0\bar{\tau} \geq r_1\tau_0 + \frac{\hat{r}_2\tau^2}{2}
\bar{\tau} = \min\left(\frac{v_0 - r_1\theta s \tau_0}{\hat{r}_2}, 1\right)
\bar{\tau} = \min\left(\frac{v_0 - r_1\theta \tau_0}{\hat{r}_2}, 1\right)
\bar{\tau} \geq 0, \tau \geq 0.
\] (66)

The problem is equivalent to the case where the value of service does not change with time as long as \(c \geq \kappa\). This suggests that all results in two-part tariff are the same except for the constant difference in the optimal value.

When \(c < \kappa\), i.e. \(\hat{c} < 0\), we can show that \(\bar{\tau}^* = 1\). Following (19) in the proof of Proposition 2, as \(\hat{c} < 0\), we obtain:
\[
\frac{d\Pi_{t|d}}{d\tau} = v_0 - r_2\bar{\tau} + (r_2 - c) \frac{\partial \bar{\tau}}{\partial \tau} \bar{\tau} \geq 0,
\]
given that we know \(v_0 - r_2\bar{\tau} \geq 0\) and \(\frac{\partial \bar{\tau}}{\partial \tau} \geq 0\). Hence, \(\bar{\tau}^* = 1\). The rest of the analysis is the same as Proposition 2.

To better understand settings where \(\kappa > c\), i.e. \(\hat{c} < 0\), we also performed numerical experiments with negative \(c\) values. This is equivalent to solving the above formulation with \(\hat{c} < 0\). We assumed \(v_0 = 10\) and the range of \(\tau_0\), \(\theta\) and \(s\) values discussed in section 4.4.1. We found that \(\bar{\tau} = 1\) and \(\bar{\tau} = 1\) in these cases. This is the same as in settings where \(c > 0\) but \(c\) is sufficiently smaller than \(v_0\), for instance \(c = 4, 6, 8\) when \(v_0 = 10\). Moreover, the conditional fixed fee and fixed fee perform as well as two-part-tariff in these settings, independent of whether \(c\) is negative or if \(c\) is significantly less than \(v_0\).