Physician Overtreatment and Undertreatment with Partial Delegation*

Dmitry Lubensky† and Eric Schmidbauer‡

November 24, 2013

Abstract

The physician induced demand literature finds that doctors tend to overtreat patients for financial gain. We analyze this phenomenon when patients are rationally skeptical of doctors’ motives and can reject a proposed treatment. We find the classic physician induced demand approach understates patients’ welfare loss: treatment on average is excessive but also less medically appropriate, and the latter effect may dominate. Inappropriate treatment arises from the doctor’s strategic misdiagnosis to forestall rejection, but this problem can be attenuated by insurance which better aligns incentives and improves communication. We resolve an open question in the partial delegation literature by showing that a generalization of the Krishna and Morgan (2001) equilibrium is the most informative equilibrium that survives the intuitive criterion in a setting that nests both our and their model.

Keywords: physician induced demand, over-utilization, non-compliance, partial delegation, cheap talk

JEL Classification: D82, I10

*We give a special thanks to Rick Harbaugh. For helpful comments we also thank Mike Baye, Hai Che, Seth Freedman, Haizhen Lin, Veronika Pool, Brandon Pope, Jeff Prince, Geoff Sprinkle, Kosali Simon, seminar participants at the BEPP Brown Bag and conference participants at the 2012 Midwest Economic Theory Conference in St. Louis, 2013 Midwest Economics Association Annual Meeting in Columbus, 2013 INFORMS Healthcare Conference in Chicago, and 2013 Marketing Science Conference in Istanbul.

†Indiana University, Kelley School of Business. dlbensky@indiana.edu.

‡Indiana University, Kelley School of Business. eschmidb@indiana.edu.
1 Introduction

Excessive treatment is a widely recognized problem in the United States health care market, with estimates of avoidable clinical care at 30% of total health care spending. An often cited cause of this phenomenon is the financial incentive of doctors to prescribe more treatment than is medically prudent (Emanuel and Fuchs, 2008). Doctors prescribe medicines with a higher markup when they profit from dispensing medicine (Iizuka, 2007), perform more cesarean deliveries when fee differentials between normal and cesarean childbirths are higher (Gruber et al., 1999), and are more likely to recommend radiation treatment for cancer when the prescription is self-referring (GAO, 2013).

The physician induced demand (PID) hypothesis posits that due to their informational advantage doctors can induce higher preferences for treatment by overstating the severity of a patient’s condition (Evans, 1974; McGuire, 2000). While this theory predicts overtreatment, its implication of a fully compliant patient is inconsistent with findings that physicians’ orders are often ignored due to a lack of trust (Brownlie et al., 2008), specifically when the patient suspects a doctor’s financial motive (Chen and Vargas-Bustamante, 2013).

When patients are rationally skeptical of doctors’ motives it is no longer clear that the PID hypothesis holds. If the doctor expects the patient to be suspicious of the diagnosis, might the doctor propose a weaker instead of stronger treatment to keep the proposed treatment from being rejected? Or might the doctor exaggerate the diagnosis even more to make the patient too scared to reject the treatment? Is the equilibrium incentive to under- or overtreat constant across all health states or are some illnesses more likely to be under- or overtreated than others? Finally, would a policy of reducing the financial incentives of doctors, such as reducing reimbursements, banning self-referrals, or instituting tort reform, improve the welfare of patients?

To address these questions, we model the interaction between an informed doctor and an uninformed patient as a game of partial delegation (e.g., Gilligan and Krehbiel, 1987; Krishna and Morgan, 2001). The doctor observes the health of the patient and prescribes a treatment, which the patient may either accept or reject in favor of

---

whatever treatment is available without the doctor’s prescription.\textsuperscript{2} Thus the doctor plays both the role of an informed expert and an authority that determines the set of available options. The doctor and patient agree on the medical prudence of a treatment, but disagree due to financial considerations. Additional treatment is more costly for the patient and is more profitable for the doctor, so that for any illness the doctor prefers more treatment than the patient.

We find that on net the patient undergoes more treatment in equilibrium than he would prefer, so the PID hypothesis holds. In fact, the average treatment is exactly what would ensue if the patient blindly followed the doctor’s orders as in the classic approach to PID without patient skepticism. More importantly, we demonstrate that the classic PID hypothesis significantly understates the welfare impact of doctors’ financial incentives by ignoring strategic considerations. Patients not only receive too much treatment but the treatment they receive is also on average less medically appropriate.

Less appropriate treatment arises from the doctor altering his diagnosis to overcome patient skepticism. For mild illnesses, the doctor prescribes a minimal treatment to avoid being rejected and for intermediate illnesses, the doctor exaggerates the diagnosis to induce the patient to accept a high treatment. This leads to under- and overtreatment in that for some illnesses the patient receives less treatment than either he or the doctor would prefer and for other illnesses he receives more treatment than either would prefer. The presence of over- and undertreatment is important for patient welfare: although it turns out not to impact the level of average treatment, it reduces the average medical appropriateness of treatment. The impact of the doctor’s financial incentive on patient welfare through the medical appropriateness of treatment is of a similar magnitude, and sometimes larger, than the impact through treatment level. We demonstrate the existence of over- and undertreatment regions is robust to model extensions in which patients may seek a second opinion or in which some consumers naively accept all treatment recommendations.

As fees for medical services increase, both overtreatment and undertreatment also increase. A higher fee increases the distance between the doctor’s and patient’s preferred treatments, which exacerbates the doctor’s incentive to strategically misdiagnose and leads to less appropriate treatment. We use this principle to investigate

\textsuperscript{2}In the main model the outside option is no treatment. In a later section we allow the patient to seek out a second opinion.
the effect of actuarially fair insurance on treatment. A patient who purchases insurance reduces his ex-post cost of medical services, bringing his incentives closer to the doctor’s and leading to more effective communication and thus more appropriate treatment. On the other hand, having insurance makes a patient less likely to reject treatment, for which he must pay ex-ante with a higher premium. We demonstrate that when the patient’s health state is distributed uniformly, the former communication effect dominates the latter moral hazard effect and the patient prefers full insurance. Thus even risk neutral patients find insurance valuable as a means to reduce the doctor’s incentive to strategically misdiagnose.

Our paper is related to the work of Pitchik and Schotter (1987) and De Jaegher and Jegers (2001) who analyze a cheap talk game in which doctors make a recommendation to skeptical patients who can obtain any available treatment. The authors depart from the standard Crawford and Sobel (1982) framework by assuming that the doctor prefers the maximal treatment regardless of the patient’s health and that doctors and patients have a different ordering over treatments. No pure strategy equilibrium can be supported and instead there exists an equilibrium in mixed strategies in which under- and overtreatment occur. Since in our model the doctor’s preferred treatment depends on the patient’s health, patients are not free to choose among all treatments, and the preference ordering adheres to the Crawford Sobel paradigm, we obtain qualitatively different predictions. For example, an increase in the patient’s out-of-pocket price of the high treatment in De Jaegher and Jegers (2001) increases patient welfare while we find the opposite in our setting. Finally, other factors affecting the doctor’s incentive to overprescribe treatment have been explored, including the role of reputation (Dranove, 1988) and second opinions (Rochaix, 1989; Pitchik and Schotter, 1993).

We model the doctor-patient relationship in a partial delegation framework since it captures the common situation where patients have veto-power over treatments but do not have the authority to unilaterally pursue their own treatment. Alternatively, communication could be modeled within a cheap talk framework where the patient has decision-making power and the physician acts solely as an advisor; e.g., when the patient discusses over-the-counter medicine with a pharmacist. In this case the Crawford and Sobel (1982) model predicts that communication is noisy so there is both over- and undertreatment, but that average treatment is no higher or lower than without communication; i.e., there is no PID overall. At the other extreme the
relationship could be modeled as full delegation (Dessein, 2002) where the patient cedes all decision-making power to the doctor; e.g., when a patient is institutionalized. In this case patient skepticism is irrelevant and the doctor always gets his preferred choice so there is always PID as in the classic approach without any variation due to strategic diagnosis. Comparing these approaches, partial delegation not only captures what is arguably the typical doctor-patient relationship, but is the only one of the three standard frameworks that features non-trivial PID.

In addition to applying partial delegation to the issue of PID, we also contribute to the general literature on partial delegation models in which a principal delegates decision making authority to an informed agent but retains the right to veto in favor of the status quo option. Of particular relevance is Krishna and Morgan’s (2001) analysis of legislative rules that models an informed but biased committee comprised of homogeneous members that makes policy proposals to an uninformed legislature.3 For a range of values of the status quo they compare the informativeness of equilibria under cheap talk versus partial delegation. While they find a partial delegation equilibrium that dominates all cheap talk equilibria, whether or not this equilibrium is the best partial delegation equilibrium has remained an open question. Our strategic environment differs from theirs in that the status quo in our model is no treatment while in their model the status quo is a moderate policy in the interior of the action space. In a general setting that nests both models, we characterize the set of equilibria which survive the intuitive criterion and identify the most informative equilibrium.4 In the process we confirm that Krishna and Morgan’s equilibrium is most informative given their status quo assumption, thereby providing additional justification for the use of their equilibrium in other contexts where it is applied.

The rest of the paper proceeds as follows. Section 2 outlines the model while Section 3 characterizes the most informative equilibrium that survives the intuitive criterion and explores comparative statics on policy variables affecting the financial incentives of doctors and patients. In Section 4 we explore the role of health insurance as a means for a patient to commit to accepting the doctor’s treatment. Section 5 confirms that over- and undertreatment persist even when patients can obtain a

\footnote{Gilligan and Krehbiel (1987; 1989) were the first to posit an informational rationale for rules that restrict amendments to committee proposals, while Krishna and Morgan followed up on their analysis.}

\footnote{Marino (2007) uses the intuitive criterion in a partial delegation setting, though the exercise is to check the criterion is satisfied by a particular equilibrium rather than characterizing the full set.}
second opinion or some proportion of patients always naively comply with a doctor’s recommendation. Section 6 concludes.

2 Model

A patient’s health state $\theta$ is distributed uniformly on $[0, 1]$, with higher values corresponding to more serious illnesses. A doctor observes the realization of $\theta$ and makes a recommendation $m \geq 0$ for treatment. The patient knows only the distribution of $\theta$ and can either take the prescribed treatment $a = m$ or a status quo $a = 0$ representing no treatment.

Payoffs for the doctor and patient are

$$u_d(a \mid \theta) = -\frac{1}{2} (\theta - a)^2 + (p - c) a,$$

$$u_p(a \mid \theta) = -\frac{1}{2} (\theta - a)^2 - pa.$$  

The first term in each payoff reflects the medical prudence of a treatment, on which both the doctor and patient agree. The second term captures financial incentives, whereby the patient pays the doctor a price $p$ for each unit of treatment and the doctor incurs a marginal cost $c$. Both $p$ and $c$ are exogenous and we assume that $p \geq c$. For any health state $\theta$ the doctor’s preferred treatment is

$$a_d^*(\theta) = \theta + (p - c),$$

and the patient’s preferred treatment is

$$a_p^*(\theta) = \theta - p.$$

Preferences closely resemble those in standard cheap talk (e.g., Crawford and Sobel, 1982) and delegation (e.g., Krishna and Morgan, 2001; Dessein, 2002) models in which the preferred action of the doctor exceeds that of the patient by a fixed amount captured by a “bias” parameter, and the loss from suboptimal actions is quadratic. By explicitly incorporating financial incentives here, we can perform comparative statics on the price and cost of treatment, and later on insurance, that would not be feasible with a standard reduced-form bias parameter.
We solve for a perfect Bayesian equilibrium, namely a strategy $m(\theta)$ for the doctor and a strategy $a(m)$ and posterior distribution $F_m(\theta)$ for the patient so that strategies are mutual best responses and beliefs are formed by Bayes rule whenever possible. For the purpose of equilibrium characterization it will be sufficient to consider the patient’s posterior expectation $\mu(m) \equiv \int \theta \, dF_m(\theta)$ conditional on message $m$.

## 3 Equilibrium and Comparative Statics

Under the strict notion of PID the doctor induces his preferred treatment for every health state. We begin this section by demonstrating this notion of PID does not survive when patients are skeptical and can reject treatment.

**Lemma 1** There is no equilibrium in which $a(m) = \theta + (p - c)$ for all $\theta$.

This follows from the fact that in such a conjectured equilibrium, the patient perfectly infers his health. Specifically, when the doctor prescribes action $p - c$, the patient understands that his health is $\theta = 0$, for which his outside option is best and thus he rejects. While the strict notion of PID does not survive, we will show it may still exist in this context if it is interpreted in terms of average treatment.

Next we characterize our equilibrium and restrict attention to parameters for which doctor and patient preferences are sufficiently aligned, namely we assume $3p - c \leq \frac{1}{2}$ throughout the paper.\(^5\)

**Proposition 1** Let $3p - c \leq \frac{1}{2}$. Then $m(\theta)$, $a(m)$, and $\mu(m)$ constitute a perfect Bayesian equilibrium when defined as follows. The doctor prescribes

\[
m(\theta) = \begin{cases} 
p - c & \text{if} \quad 0 \leq \theta < 3p - c \\
7p - 3c & \text{if} \quad 3p - c \leq \theta \leq 6p - 2c \\
\theta + p - c & \text{if} \quad 6p - 2c < \theta \leq 1
\end{cases}
\]

\(^5\)Parameters outside this range can still support equilibria with information transmission but lead to corner solutions that complicate the analysis and are omitted for clarity. In Appendix A, we remove this parameter restriction when solving for the full set of equilibria that survive the intuitive criterion.
The patient accepts treatment if and only if prescribed as above and has beliefs

$$
\mu(m) = \begin{cases} 
0 & \text{if } 0 \leq m < p - c \\
\frac{3p-c}{2} & \text{if } m = p - c \\
\frac{m-(p-c)}{2} & \text{if } p - c < m < 7p - 3c \\
\frac{9p-3c}{2} & \text{if } m = 7p - 3c \\
m - (p - c) & \text{if } 7p - 3c < m 
\end{cases}
$$

The proof that this constitutes a perfect Bayesian equilibrium is a special case of the more general equilibrium characterization in Appendix A and therefore omitted here. An example of the equilibrium for particular values of $p$ and $c$ is depicted in Figure 1. The two dashed lines represent preferred actions: the line below represents the patient’s preferences and the line above the doctor’s. For mild illnesses the patient receives a single minimal treatment that is less than what the doctor prefers and, for some states, also less than what the patient prefers. For intermediate illnesses, the patient receives a large treatment that is higher than his and even the doctor’s preferred action. For sufficiently serious illnesses, the patient receives the treatment preferred by the doctor. Treatments in the intervals $a \in (0, p - c) \cup (p - c, 7p - 3c)$ are off-the-equilibrium path and thus admit a variety of beliefs. The particular beliefs above were chosen for concreteness and we will later show they satisfy the intuitive criterion.

The logic behind the equilibrium reflects the fact that the patient has the option of no treatment and is skeptical of the doctor’s motives. The patient accepts a significant treatment only when he infers from the diagnosis that he is quite sick. This accounts for the fact that in equilibrium intermediate treatments $a \in (p - c, 7p - 3c)$ are rejected. When the doctor’s preferred treatment falls in this range, he must choose whether to overstate beyond his financial incentive or conversely to understate and induce a minimal treatment instead of being rejected altogether. In equilibrium the doctor does both: he induces a single minimal treatment for relatively healthier states ($0 \leq \theta \leq 3p - c$) and a single overly strong treatment for relatively sicker states ($3p - c \leq \theta \leq 6p - 2c$). The doctor’s inability to customize his diagnosis based on his information reduces average treatment appropriateness, and in this way patient’s skepticism exacerbates the effect of the doctor’s financial incentives.
Figure 1: In solid, equilibrium treatment when $p = 0.1$ and $c = 0.05$. In dark dashed, the patient’s preferred treatment as a function of the health state. In light dashed, the doctor’s preferred treatment as a function of the health state.

**Equilibrium Selection**

It is well-known that in games of communication, and in particular in our partial delegation setting, there is a multiplicity of equilibria. For example, there exists a trivial equilibrium in which the patient accepts only zero treatment, which is supported by off the path beliefs that any prescription for a strictly positive treatment is made when the doctor observes $\theta = 0$. While this particular equilibrium does not survive standard refinements, the equilibrium set contains other non-trivial equilibria which may not be ruled out so easily.

The issue of equilibrium selection has played an important role in this literature, namely related to the question of whether partial delegation or cheap talk is the better communication mechanism. Gilligan and Krehbiel (1987, 1989) identify a partial delegation equilibrium and a cheap talk equilibrium and demonstrate that depending on parameter values either can dominate the other. Krishna and Morgan (2001) follow up on this work by finding a different partial delegation equilibrium that dominates all cheap talk equilibria, thus concluding that partial delegation is the better protocol. However, the authors leave open whether the equilibrium they find is most informative. Since the aim of our work is not to compare mechanisms
but rather to explore equilibrium properties, the unresolved question of selection in the partial delegation framework is central to our analysis.

Our approach is to focus only on equilibria that survive the Cho and Kreps (1987) intuitive criterion, and of this set select the equilibrium which is most informative, defined as follows:

**Definition 1**  
*The informativeness of an equilibrium is* \( \text{Var} [\theta - a(\theta)] \).  

Krishna and Morgan provide justification for informativeness as an appropriate metric by which to measure equilibria. In our model and the Crawford and Sobel framework in general, the quadratic loss function implies that the preferences of the doctor and the patient can be decomposed into preferences over the average treatment and the variance of treatment. Conditional on average treatment, both the doctor and patient are better off when variance is reduced, thus the set of treatment allocations which minimize variance can be thought of as a Pareto frontier. Under cheap talk, all equilibria are associated with the same expected treatment, hence if welfare is the weighted sum of the utilities of the doctor and the patient, maximizing informativeness is equivalent to maximizing welfare regardless of the weights. In the partial delegation environment, different equilibria may be associated with different expected actions, thus moving to an equilibrium with less variance but a different expected treatment may harm either the doctor or the patient. However, the move to less variance is a move toward the Pareto frontier, and since in our specification preferences are quasilinear in health and money, this move would make both parties better off given an appropriate cash transfer.\(^6\)

Using both the intuitive criterion and informativeness for equilibrium selection, we obtain the following results.

**Proposition 2**  
The equilibrium in Proposition 1 is the most informative equilibrium that survives the intuitive criterion.

**Proposition 3**  
The closed rule equilibrium in Proposition 8 in Krishna and Morgan (2001) is the most informative equilibrium that survives the intuitive criterion in their parameter range.

---

\(^6\) An alternative approach is to rank equilibria based on a specific weighted average of the utilities of the doctor and patient. The choice of such weights would be arbitrary, especially given that both parties have some control over the treatment decision.
The proof of both of these propositions is obtained by characterizing the most informative equilibrium that survives the intuitive criterion in a more general setting which nests both our model and the model of Krishna and Morgan (2001). In particular, the framework allows for a fuller range of treatment preference parameters and a status quo of zero as in our model or interior status quo as in Krishna and Morgan (2001). The details can be found in Appendix A.

**Equilibrium Properties**

Lemma 1 demonstrates that the strict notion of the PID hypothesis, in which the doctor induces his preferred action for every illness, does not survive in our setting with skeptical patients. However, the following corollary demonstrates that PID still holds in our context when reinterpreted in terms of average treatment.

**Corollary 1** *In the equilibrium in Proposition 1, $E[a(\theta)] = E[\theta] + (p - c)$.*

Stated differently, the expected treatment in equilibrium is exactly what would ensue if the doctor could induce his preferred treatment for every illness. An empirical analysis which finds an increased average treatment level in the presence of doctors’ financial incentives is thus consistent both with the classic notion of PID and with our partial delegation framework. However, due to the information loss associated with partial delegation, the classic PID model understates the welfare effect of doctor’s financial incentives, and we explore this here in more detail.

**Definition 2** *State $\theta$ is overtreated if $a(\theta) > a_d^*(\theta)$ and undertreated if $a(\theta) < a_p^*(\theta)$.*

Thus an illness is overtreated if the treatment exceeds that preferred by the doctor and undertreated when it falls below that preferred by the patient. We use weak definitions of over- and undertreatment by remaining agnostic about the relative weights on the utilities of the doctor and patient. Alternatively, one may construct a sharper definition of over- and undertreatment by defining a benchmark that would be chosen by a social planner. For example, a social planner that weighs the medical prudence of an action against its true cost would have a utility $u_{sp}(a|\theta) = -\frac{1}{2}(\theta - a)^2 - ca$ and an optimal treatment plan $a_{sp}^*(\theta) = \theta - c$. It is clear that the existence of over- and undertreatment under our weak definition implies the existence of over- and undertreatment with respect to this benchmark, in fact with respect to any
benchmark in which the preferred action falls between the preferred actions of the doctor and patient.

In Figure 1, undertreatment occurs for some states for which the minimal treatment \((a = p - c)\) is undertaken and overtreatment occurs for all states for which the high treatment \((a = 7p - 3c)\) is undertaken. More generally, we show that over- and undertreatment are guaranteed to occur for all parameters \(p\) and \(c\) admissible in Proposition 1.

**Corollary 2** In the equilibrium in Proposition 1 both under- and overtreatment exist. Namely, states \(\theta \in (2p - c, 3p - c]\) are undertreated and \(\theta \in (3p - c, 6p - 2c)\) are overtreated.

**Proof** By construction of the equilibrium in Proposition 1, it can be verified that for health states near but below the state \(\theta = 3p - c\) at which there is a discontinuity in treatment, the equilibrium treatment of \(p - c\) is below that preferred by the patient. For all states above \(\theta \in [3p - c, 6p - 2c)\) the treatment is greater than that preferred by the doctor. ■

The notions of over- and undertreatment help us understand the informational effects of a doctor’s financial incentives beyond their effect on average treatment level. For any two treatment plans \(a_1(\theta)\) and \(a_2(\theta)\) for which \(E[a_1(\theta)] = E[a_2(\theta)]\), both the doctor and the patient prefer plan \(a_1(\theta)\) whenever it has less over- and undertreatment than \(a_2(\theta)\). This is because plan \(a_1(\theta)\) induces actions that are closer to those preferred by both the doctor and patient than treatment plan \(a_2(\theta)\), and thus we refer to it as more appropriate.

The connection between over- and undertreatment and patient welfare can be formally established by decomposing the effect of doctors’ financial incentives on equilibrium treatment. First, denote by \(U_p(a(\theta)) \equiv E[u_p(a(\theta)|\theta)]\) the patient’s expected utility from following treatment plan \(a(\theta)\). Recall that \(a^*_p(\theta)\) and \(a^*_d(\theta)\) are the patient’s and doctor’s first best treatment plans, and let \(a_{eqm}(\theta)\) be the equilibrium treatment plan. The equilibrium utility loss for the patient relative to his first best can be expressed as

\[
\begin{align*}
\text{Total Welfare Loss} & = U_p(a^*_p) - U_p(a_{eqm}) \\
\text{Loss from Classic PID} & = U_p(a^*_p) - U_p(a^*_d) \\
\text{Loss from Treatment Appropriateness} & = U_p(a^*_d) - U_p(a_{eqm})
\end{align*}
\]
The first term on the right hand side describes the difference in the patient’s utility between obtaining his or the doctor’s first-best treatments. We label this term the classic PID effect since Corollary 1 established that the average equilibrium treatment is the same as if the doctor induced his preferred treatment for every state. Because it is never optimal for the doctor to induce over- or undertreatment as we have defined them, the impact of these outcomes must be contained in the second term of the decomposition, which we now demonstrate is strictly positive.

**Corollary 3** The loss in treatment appropriateness is strictly positive; that is,
\[ U_p(a^*_p) - U_p(a_{eqm}) > 0. \]

**Proof** The patient’s expected utility from any treatment profile \( a(\theta) \) can be expressed as the sum of two terms:
\[ U_p(a(\theta)) = -\frac{1}{2} E[(a(\theta) - \theta)^2] - pE[a(\theta)]. \]

Moving from the doctor’s first-best treatment profile to the equilibrium treatment profile constitutes a mean preserving spread in \( a(\theta) - \theta \). Thus for equilibrium treatments the first term is larger than and the second term is equal to that under the doctor’s first-best treatment profile. ■

Corollary 3 confirms that by ignoring treatment appropriateness classic PID understates the impact of a doctor’s financial incentives on patient welfare. In fact, we now show that treatment appropriateness can be a more important consideration than treatment level.

**Corollary 4** When the marginal cost of treatment \( c \) is high, treatment inappropriateness harms patient welfare more than treatment level; that is,
\[ U_p(a^*_p) - U_p(a^*_d) < U_p(a^*_d) - U_p(a_{eqm}). \]

The proof is included in Appendix B, however here we provide a brief sketch. Conjecture a situation in which the doctor fully passes through an increase in cost \( c \). Since the doctor’s markup \( p - c \) is unchanged, his preferred treatment remains unchanged and by extension so does the average treatment undertaken in equilibrium. However, since the patient’s price of treatment increases, the incentives of the doctor and patient diverge and this leads to less appropriate treatment. When \( c \) is increased
sufficiently in this margin-neutral way, the loss from ineffective treatment can surpass the loss from the treatment level.

Our results about the expected level of treatment in Corollary 1 and the expected appropriateness of treatment in Corollary 3 can be understood in the context of how the partial delegation environment is situated between pure cheap talk (Crawford Sobel, 1982) and full delegation (Dessein, 2002). With cheap talk, the patient is unrestricted in his choice of treatment and for every diagnosis maximizes expected utility conditional on his posterior. Because the patient’s posterior is correct, it is as if the patient chooses his first best treatment on average, thus there is no PID, but because communication is coarse the patient sometimes chooses ineffective treatment.\footnote{This argument follows formally when optimal treatment is linear in patient beliefs, as it is in the quadratic loss specification.}

The other extreme is full delegation, in which the patient has no choice and must accept the doctor’s preferred action. There is no noise, which causes treatment to be appropriate, and PID exists in the strict sense. Under partial delegation, the doctor is neither simply an advisor nor the decision maker, instead he makes a take-it-or-leave-it offer. Again, the patient’s posterior is correct in equilibrium but his optimal treatment is not available. Instead, the doctor constructs the offer so as to make the patient indifferent between accepting and rejecting, extracting the surplus from the interaction and on average inducing his preferred treatment.\footnote{The analogy holds for mild and intermediate illnesses; for severe illnesses the doctor can induce his preferred treatment without having to make the patient indifferent to rejecting.}

With partial delegation, PID exists in terms of average treatment and noisy communication reduces treatment effectiveness. Thus partial delegation is the only communication protocol in which the doctor’s financial incentives affect treatment both through average level and average effectiveness.

**Comparative Statics**

Our model has two parameters, $p$ and $c$, that correspond to policy interventions such as changes in reimbursement rates, subsidies for medical supplies, changes in medical technology, etc. Here we explore the impact of such interventions on the utilities of the doctor and the patient. In addition, we track the impact on the total level of over- and undertreatment, which we define as the ex-ante probability that a state
with over- or undertreatment occurs.\textsuperscript{9}

**Proposition 4** The following table summarizes the effect of \( p \) and \( c \) on patient and doctor utility, overtreatment, and undertreatment.

<table>
<thead>
<tr>
<th>Increase in</th>
<th>( U_{\text{patient}} )</th>
<th>( U_{\text{doctor}} )</th>
<th>Overtreatment</th>
<th>Undertreatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p )</td>
<td>(-)</td>
<td>(+/-)</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>( c )</td>
<td>(+)</td>
<td>(-)</td>
<td>(-)</td>
<td>(0)</td>
</tr>
<tr>
<td>( p ) and ( c )</td>
<td>(-)</td>
<td>(-)</td>
<td>(+)</td>
<td>(+)</td>
</tr>
</tbody>
</table>

These results are derived in Appendix B by using the closed form solution of the equilibrium in Proposition 1. Consider first the table’s top two rows. An increase in \( p \) affects patient utility in three ways. First, its direct effect is to raise the patient’s total health expenditure if treatment remains unchanged. It also has an indirect PID effect whereby a higher price causes the doctor to induce higher average treatment. Finally, an increase in \( p \) exacerbates the difference in doctor and patient preferences, thus causing less appropriate treatment through more over- and undertreatment. The net effect of a higher \( p \) on patient welfare is thus unambiguously negative. The effect of \( p \) on the utility of the doctor can similarly be decomposed into the three separate effects. The direct effect is positive while the PID effect is zero because the doctor always obtains his preferred average treatment in equilibrium. Because treatment effectiveness declines, the net effect on the doctor’s utility is ambiguous.

A similar logic applies to the effect of \( c \) on the utilities of the doctor and patient. A seemingly surprising result emerges: a higher cost \( c \) of treatment increases the welfare of the patient. This follows within the model since a higher \( c \) corresponds to closer doctor and patient preferences. However, this particular comparative static assumes none of the cost increase is passed through to patients. In the other extreme, costs can be fully passed through as considered in the third row of the results table and in Figure 2. Recall that a higher \( p \) decreases treatment effectiveness while a higher \( c \) increases it. As seen in the table, the net effect of an increase in \( p \) on over- and undertreatment is stronger than the countervailing effect of an (equal magnitude) increase in \( c \). This stems from the fact that when a cost increase is fully passed

\textsuperscript{9}An alternative definition is the patient’s utility loss relative to his first best in regions of undertreatment and the doctor’s utility loss relative to his first best in regions of over-treatment. In our equilibrium this definition induces the same ranking of treatment plans and we use ours for simplicity.
through the doctor’s profit margin and hence preferences remain unchanged while the patient faces a higher price and so prefers less treatment. On net, preferences diverge further so that treatment is less appropriate and the expected utilities of the doctor and patient decline.

4 Insurance: Moral Hazard and Communication

While doctors’ financial incentives are often cited as a cause of rising health care costs, another commonly identified culprit is that most patients pay for treatment with the aid of health insurance (DOJ and FTC report, 2004). Although insurance companies actively manage coverage to deny unnecessary medical procedures, much of the decision-making power still rests with the patient, who faces a fraction of the full charge for a medical service and may not fully internalize its cost (van Dijk et al., 2013). The patient may thereby undergo unnecessary procedures, i.e., the moral hazard of insurance.

On other hand, in light of the results of the previous section, a patient that pays only a fraction of the treatment cost has incentives that are more closely aligned with those of the doctor, which should result in more appropriate treatment on average. In the ensuing analysis, we explore the trade-off between moral hazard and treatment effectiveness and demonstrate that treatment effectiveness dominates.
We augment the model by introducing an ex-ante stage, at which a patient purchases an insurance contract. The contract specifies the patient’s coinsurance rate \( \gamma \in [0, 1] \) and premium \( P(\gamma) \), so that the patient’s expected utility is

\[
    u_p(\gamma) = -\frac{1}{2}(\theta - a(\theta))^2 - \gamma pa - P(\gamma).
\]

When a policy with coinsurance rate \( \gamma \) has been purchased the premium \( P(\gamma) \) is sunk and the patient prefers a weakly higher treatment for all health states than without insurance:

\[
    a_p^*(\theta) = \theta - \gamma p.
\]

We assume that the insurance market is competitive and \( P(\gamma) \) is actuarially fair, thus

\[
    P(\gamma) = (1 - \gamma)pE[a(\theta | \gamma)].
\]

Conditional on a treatment plan insurance does not save the patient money: his ex-post payment is reduced and instead shifted to a sunk payment up-front. However, the purchase of insurance commits the patient to be more likely to accept treatment, thereby potentially changing the amount and effectiveness of treatment he receives in equilibrium. To formalize this, we characterize an equilibrium of the doctor and patient interaction with insurance:

**Proposition 5** Let \( 3p - c \leq \frac{1}{2} \) and \( \gamma \in [0, 1] \). Then \( m(\theta) \), \( a(m) \), and \( \mu(m) \) constitute a perfect Bayesian equilibrium when defined as follows. The doctor prescribes

\[
    m(\theta) = \begin{cases} 
    p - c & \text{if } 0 \leq \theta < (1 + 2\gamma)p - c \\
    (3 + 4\gamma)p - 3c & \text{if } (1 + 2\gamma)p - c \leq \theta \leq (2 + 4\gamma)p - 2c \\
    \theta + p - c & \text{if } (2 + 4\gamma)p - 2c < \theta \leq 1
    \end{cases}
\]
The patient accepts treatment if and only if prescribed as above and has beliefs

\[
\mu(m) = \begin{cases} 
0 & \text{if } 0 \leq m < p - c \\
\frac{(1+2\gamma)p-c}{2} & \text{if } m = p - c \\
\frac{m-(p-c)}{2} & \text{if } p - c < m < (3 + 4\gamma)p - 3c \\
\frac{3(1+2\gamma)p-3c}{2} & \text{if } m = (3 + 4\gamma)p - 3c \\
m - (p - c) & \text{if } (3 + 4\gamma)p - 3c < m
\end{cases}
\]

Note that the equilibrium described in the previous section is a special case of the equilibrium with insurance in which \( \gamma = 1 \). Figure 3 illustrates the effect of more insurance (lower \( \gamma \)). Region A depicts health states for which treatment increases as the result of added insurance and thus corresponds to the moral hazard effect. Region B depicts health states for which the patient undergoes less treatment when buying more insurance. This effect arises because a patient with lower coinsurance is less likely to reject, and consequently the doctor’s incentive to overstate his diagnosis is reduced. Figure 3 thus demonstrates that the moral hazard effect on average treatment is offset by the reduction in overtreatment stemming from a closer alignment of doctor and patient incentives.

The proof of Proposition 6 below shows that the expected level of equilibrium treatment is independent of \( \gamma \). In fact, as in the previous section the average treatment remains \( E[\theta] + (p - c) \), that which would ensue if the doctor induced his first-best treatment for every illness. This demonstrates that an increase in the coinsurance \( \gamma \) rate merely reallocates health expenditures from ex-ante to ex-post, while keeping the sum of these two expenditures constant. This implies that the effect on doctor and patient welfare is entirely through the effect on communication, and communication is best when doctor and patient incentives are closest, i.e., the patient is fully insured.

**Proposition 6** Full insurance (i.e., \( \gamma = 0 \)) is preferred by both the doctor and patient.

The proof can be found in Appendix B. In contrast to a full information context in which one would find insurance increases the medical action taken for every health state, we find it increases treatment for states that were previously undertreated and
Figure 3: A decrease in $\gamma$ (more insurance) increases expected treatment by the area of region A but decreases it by the area of region B. For $\theta \sim U[0, 1]$ these two exactly balance so that expected treatment is constant.

reduces treatment for previously overtreated states. Therefore the gains in patient utility derive from a better matching of treatments to health states while total health expenditures are constant. For his part, the doctor’s expected profits from treatment are constant in the coinsurance requirement though he too gains from the better matching of treatments to health states.$^{10}$

5 Extensions

Naive Patients

We have demonstrated that when all patients are rationally skeptical, doctors’ financial incentives lead to over- and undertreatment which hurt both the doctor and the patient. On the other hand, the classic PID model may be interpreted as one in which all patients are naive in blindly following the doctor’s orders. As a robustness check, in this section we allow for a proportion of naive patients and explore whether over- and undertreatment remain.

$^{10}$Our result is reminiscent of Dessein (2002) in which he investigates whether a principle could gain from using an intermediary with a given bias of his choosing.
**Proposition 7** For a sufficiently small proportion of naive patients, the equilibrium in Proposition 1 survives. Consequently there exists over- and undertreatment.

See Appendix B for the proof. The existence of naive patients gives the doctor a new treatment option. For instance, the doctor that was indifferent between inducing the minimally accepted treatment \( p - c \) and the inflated treatment \( 7p - 3c \) now has the option of inducing his preferred treatment between the two but only from naive patients. When the number of naive patients is small, this option is strictly worse than the other two, thus his incentives do not change qualitatively.

When there is a large enough number of naive patients, the doctor may eventually deviate and induce his most preferred treatment from the naives while being rejected by the rational patients. Solving for the equilibrium set when the number of naive patients is large is tedious and beyond the scope of this work. However, we can quite easily obtain the following result.

**Proposition 8** When there is a positive measure of rational patients, there exists no separating equilibrium.

**Proof** Conjecture a separating equilibrium and note that since all equilibria are monotone in induced treatment (Lemma 2, Appendix A), there must exist a largest type \( \hat{\theta} \) such that all treatments proposed by types \( \theta < \hat{\theta} \) are rejected by the \( 1 - \alpha \) rational patients. Then, a type just below \( \hat{\theta} \) can misreport a slightly higher illness and in doing so reduce his utility from treating naive patients marginally while increasing his utility from now treating rational patients discretely. ■

The proposition above guarantees pooling regions, which we believe ought to lead to regions of over- and undertreatment as in our main model.

It is also interesting to consider the limiting case in which all patients are naive, as would be under classic PID. Here the equilibrium is fully separating and there is no information loss. Compared to the equilibrium in Proposition 1, both outcomes have the same expected treatment however the equilibrium with all naive patients has more appropriate treatment and is therefore preferred by the doctor and patient. This finding is similar to Dessein (2002) who demonstrates that full delegation is superior to Krishna and Morgan’s (2001) closed rule model for low levels of bias.
Second Opinions

Our model has assumed thus far that the patient’s only alternative to a doctor’s recommendation is of taking no action at all. In reality, patients often have the option to consult a second or third doctor for an additional opinion and alternative treatment options. For simplicity and to retain focus on the communication game we do not consider doctors with heterogeneous levels of expertise but rather the possible effect of competition between homogenous doctors on doctor-patient communication.

If the patient can visit multiple doctors costlessly then a fully separating equilibrium exists. Each doctor recommends the patient’s preferred action while the patient visits every doctor and accepts treatment only if all recommendations are the same. If recommendations diverge the patient believes the true health state is zero and rejects. From the perspective of the doctor, if his competitors send fully separating recommendations he gains nothing by misrepresenting his information, as this induces the patient immediately to reject.

It is more realistic however that getting a second opinion has some positive cost. Then a fully separating equilibrium in which each doctor recommends the patient’s preferred action cannot exist, since if it did no patient would incur search costs to obtain a second opinion. But given no patient searches, each doctor has incentive to deviate for every state. Instead, there exists an equilibrium in which all doctors prescribe as in Proposition 1. The patient does not search since his perceived return to doing so is zero, and instead also follows his strategy in Proposition 1.

The no-search result is an application of the “Diamond paradox” which states there is no equilibrium in which search occurs when the cost of search is positive (Diamond, 1971). The existence of a search cost brings about a hold-up problem that allows each doctor to behave as a monopolist: a doctor offering a second opinion can infer what message the first doctor sent and undercut it by some small amount $\varepsilon > 0$ less than the search cost. Anticipating this, the patient does not search in the first place. Note that in a different setting Krishna and Morgan (2001b) reach the similar conclusion that no additional information can be gained by visiting a second expert, while Pitchik and Schotter (1993) find doctor opportunism can be competed away only if a sufficiently large proportion of consumers have no search costs.
6 Conclusion

The classic theory of physician induced demand posits that financially incentivized doctors overprescribe medical services to patients who accept the doctor’s orders. When allowing for patients to be rationally skeptical of the doctor’s motives we discover the classic PID approach understates the negative welfare effects of the doctors’ incentives. Though patients are skeptical, in equilibrium doctors still induce the same average treatment as they would if patients blindly accepted the doctors’ first-best treatment, thus confirming that PID still exists in our context. However, we also discover an informational effect in which the prescribed treatment is on average less appropriate, stemming from doctors strategically misdiagnosing to convince patients to accept treatment. We find that the informational effect is of a similar magnitude as the classic PID effect, and sometimes may be the dominant effect.

We use our framework to conclude health insurance reduces the incidence and severity of over- and undertreatment. It is well-understood that the presence of health insurance can induce moral hazard in that a patient who pays only a fraction of his medical bill is likely to obtain more treatment. However, we uncover a countervailing effect in which the presence of insurance increases the appropriateness of treatment by more closely aligning the incentives of the doctor and the patient. In our model this new effect completely dominates the moral hazard effect, so that in equilibrium when the patient purchases more insurance his average treatment remains fixed while its effectiveness increases.

The impact of financial incentives on treatment appropriateness can be expressed in terms of over- and undertreatment. Both of these phenomena persist for alternate model specifications in which some patients are naive while others are rationally skeptical, or in which the patient can obtain a second opinion at a small cost.

In addition to highlighting that the classic PID model understates the effect of the doctor’s financial incentives, we also contribute to literature on partial delegation. In a similar setting to ours, Krishna and Morgan (2001) describe an equilibrium which dominates all equilibria in the Crawford and Sobel (1982) framework but leave open whether the equilibrium they find is most informative in their environment. We build on their work by applying the Cho Kreps (1987) intuitive criterion and characterizing the equilibrium set in a more general framework that nests both our and their model. We confirm their equilibrium is in fact the most informative in their parameter space.
Appendix A: Existence and Informativeness Proofs

For the sake of generality and to facilitate comparison with Krishna and Morgan (2001) and Gilligan Krehbiel (1989), we use the following model. Let the utilities of the doctor and the patient be

\[ u_d(a|\theta) = -\frac{1}{2}(\theta - a)^2 + b_d a \]
\[ u_p(a|\theta) = -\frac{1}{2}(\theta - a)^2 - b_p a \]

The doctor’s preferred treatment is \( a_d^*(\theta) = \theta + b_d \) and the patient’s preferred treatment is \( a_p^*(\theta) = \theta - b_p \), in which \( b_d \) and \( b_p \) are both weakly positive. The patient’s outside option treatment is \( 0 \leq s < 1 - b_d \), so that the outside option is guaranteed to be lower than the doctor’s maximal preferred treatment.

Note that our model is a special case in which \( b_d = p - c, b_p = p, \) and \( s = 0 \) and the Gilligan Kriehbel (1989) and Krishna Morgan (2001) model is also a special case in which \( b_d = x_c, b_p = 0, \) and \( 3x_c < s < 1 - x_c. \)

In what follows we establish several lemmas, characterize the set of equilibria, and then use informativeness to select a particular equilibrium. In doing so we provide a proof of Propositions 1, 2, and 3.

Preliminary Results

Lemma 2 Equilibrium treatments weakly increase in the health state \( \theta \).

Proof This follows from a single crossing property argument. Toward a contradiction there exist states \( \theta \) and \( \theta' \) such that \( \theta < \theta' \) and in equilibrium \( a(\theta) > a(\theta') \). If the doctor is best responding in state \( \theta \), then \( u_d(a(\theta), \theta) \geq u_d(a(\theta'), \theta) \) and his ideal treatment is weakly closer to \( a(\theta) \) than to \( a(\theta') \). However, in state \( \theta' > \theta \), the doctor’s ideal treatment is even higher which implies \( u_d(a(\theta'), \theta') < u_d(a(\theta), \theta') \), and thus leads to a contradiction. ■

Lemma 3 Suppose treatment \( a' \) is pooled on \( (\theta, \bar{\theta}) \) and \( a_d^*(\bar{\theta}) > a' \). Then there is a discrete jump to the next induced treatment \( a'' = 2(\bar{\theta} + b_d) - a' \) and it is pooled at least out to \( \hat{\theta} \equiv \bar{\theta} + 2(b_d + b_p) - a' + s \).
Proof By continuity of the doctor’s utility function, the type $\bar{\theta}$ doctor must be indifferent between inducing treatment $a'$ and $a''$; that is,

$$a'' - a^*_d(\bar{\theta}) = a^*_d(\bar{\theta}) - a' \iff a'' = 2(\bar{\theta} + b_d) - a'.$$

The treatment $a''$ is pooled because types $\theta > \bar{\theta}$ in the neighborhood of $\bar{\theta}$ prefer a lower treatment than $a''$ but treatments $a \in (a',a'')$ are not inducible.

In order for patients to accept $a''$, their posterior $\mu$ must be sufficiently high. That is, the minimal $\widehat{\theta}$ must satisfy

$$a'' - (\mu - b_p) = \mu - b_p - s \iff \widehat{\theta} = \bar{\theta} + 2(b_d + b_p) - a' + s.$$

These first two Lemmas demonstrate that all equilibria have some regularities, and we later use these results to classify equilibria into a manageable number of categories. The following two lemmas are intended to reduce later work by performing commonly recurring calculations in the ensuing proof.

Lemma 4 The patient accepts $a' > s$ on $[\theta, \bar{\theta}]$ only if $a' + s + 2b_p \leq \theta + \bar{\theta}$.

Proof If the patient’s preferred treatment $\mu - b_p$ is above $a'$ then $a' > s$ is accepted. If $a' > \mu - b_p > s$ then, due to the quadratic loss utility, $a'$ is rejected if it is farther away from $\mu - b_p$ than is the outside option $s$. That is

$$a' - (\mu - b_p) \leq (\mu - b_p) - s \iff a' + s + 2b_p \leq \theta + \bar{\theta}.$$

If $a' > s \geq \mu - b_p$ then $a'$ is rejected.

Lemma 5 Suppose treatments $a_x$ and $a_y$ are accepted in equilibrium and all treatments $a' \in (a_x,a_y)$ are rejected. For any such off-the path treatment $a'$, the lowest beliefs admitted by the intuitive criterion are $\mu = \frac{a' + a_x}{2} - b_d$ and the deviant message is rejected if and only if $a_x \leq 2(b_d + b_p) + s$.

Proof The lowest permissible belief is $\mu = (a_d^*)^{-1}\left(\frac{a' + a_x}{2}\right) = \frac{a' + a_x}{2} - b_d$, as lower types prefer treatment $a_x$, which gives them at least their equilibrium payoff, to treatment.
In order to reject a deviant message, conditional on the patient’s belief the distance from his preferred treatment to \( a' \) must be greater than the distance to the outside option \( s \):

\[
a' - (\mu - b_p) \geq (\mu - b_p) - s \iff a_x \leq 2 (b_d + b_p) + s.
\]

**Corollary 5** Let \( a \) be the highest treatment induced in equilibrium such that there is a jump to the next highest induced treatment. If \( a > 2 (b_d + b_p) + s \) then the equilibrium fails the intuitive criterion. Conversely, if \( a \leq 2 (b_d + b_p) + s \) then beliefs consistent with the intuitive criterion can support the equilibrium.

**Equilibrium Classification**

We classify equilibria into two categories: those in which the patient sometimes takes his outside option \( s \) and those in which he does not. We will show that depending on parameters the most informative equilibrium can be found in either category. To simplify the presentation of the equilibrium characterization, within each category we focus only on equilibria with the smallest pooling regions. We subsequently show that within a category, an equilibrium with smaller pooling regions is more informative than an equilibrium with larger pooling regions.

**Equilibria in which treatment \( s \) is not induced**

We decompose all equilibria in which \( s \) is not induced into subclasses of equilibria, each subclass corresponding to a lowest induced treatment \( a_1 \). We demonstrate that once \( a_1 \) is chosen, there is little freedom in the remaining construction of the equilibrium.

**Proposition 9** Let \( a_1 \) be the lowest induced treatment. In any equilibrium that satisfies the intuitive criterion in which treatment \( s \) is not induced, the doctor’s message function is of one of the two following forms:

i. If \( a_1 \in (s, b_d) \) then the doctor prescribes

\[
m(\theta) = \begin{cases} 
a_1 & \theta \in [0, b_d + s + 2b_p] \\
4(b_d + b_p) + 2s - a_1 & \theta \in (b_d + s + 2b_p, 3b_d + 4b_p + 2s - a_1) \\
\theta + b_d & \theta \in [3b_d + 4b_p + 2s - a_1, 1]
\end{cases}
\]
ii. If $a_1 \in [b_d, 2(b_d + b_p) + s]$ then the doctor prescribes

$$m(\theta) = \begin{cases} 
  a_1 & \theta \in [0, a_1 + s + 2b_p] \\
  a_1 + 2s + 2b_d + 4b_p & \theta \in (a_1 + s + 2b_p, a_1 + 2s + b_d + 4b_p) \\
  \theta + b_d & \theta \in [a_1 + 2s + b_d + 4b_p, 1]
\end{cases}$$

and all such messages are accepted.

**Proof** First, in this type of equilibrium the doctor’s preferred treatment is above $s$ for every illness, that is $b_d > s$. Otherwise, there would be at least one $\theta$ for which the doctor’s preferred treatment is the patient’s outside option, and in equilibrium this treatment is induced.

Next, consider the lowest induced treatment $a_1$ and note that $a_1 > s$. This follows because $a_1 \neq s$ by hypothesis and if $a_1 < s$ then any type that induces $a_1$ would deviate to inducing treatment $s$. By Lemma 2, $a_1$ must be induced on some interval $\theta \in [0, \theta_1)$, and by Lemma 4, $\theta_1 \geq a_1 + s + 2b_p$ in order for the patient to accept. We define $\theta_1 = a_1 + s + 2b_p$ as the minimal right endpoint necessary to induce acceptance of $a_1$.

By inspection $a_d^{\ast}(\theta_1) > a_1$, meaning that the highest type to induce $a_1$ strictly prefers a higher treatment. By Lemma 3 this implies at $\theta_1$ there is a jump in treatment
up to a specific treatment $a_2 = 2(\theta_1 + b_d) - a_1$, and that treatment at $a_2$ is also pooled on an interval $(\theta_1, \theta_2]$.

The endpoint $\theta_2$ plays an important role. First, note that the treatment $a_2$ is higher than the doctor’s preferred treatment in the neighborhood of $\theta_1$, in fact the doctor must induce $a_2$ for all states $\theta > \theta_1$ for which his preferred treatment is lower, i.e. $\theta_2 \geq (a_2^*)^{-1}(a_2)$. The question then is whether $\theta_2 = (a_2^*)^{-1}(a_2)$ is sufficiently large to induce the patient to accept, or whether it must be made larger. We now argue that whenever $\theta_2$ needs to be larger, the equilibrium fails the intuitive criterion.

If $\theta_2 > (a_2^*)^{-1}(a_2)$ then there must be a jump above $a_2$ to the next highest induced treatment. But by Corollary 5 the intuitive criterion is only satisfied if $a_2$ is sufficiently small. Note that

$$a_2 = 2(\theta_1 + b_d) - a_1 \geq 2((a_1 + s + 2b_p) + b_d) - a_1$$

$$= a_1 + 2s + 4b_p + 2b_d > 2(b_d + b_p) + s$$

Thus, since $a_2 > 2(b_d + b_p) + s$, a jump above $a_2$ means the intuitive criterion is not satisfied. By Lemma 3, $\theta_2 = \theta_1 + 2(b_d + b_p) - a_1 + s$ is the lowest endpoint of this pooling region for which the patient’s posterior is sufficiently high to accept. To ensure no jump occurs, it is necessary that

$$(a_2^*)^{-1}(a_2) \geq \theta_2$$

$$\iff \quad 2\theta_1 + b_d - a_1 \geq \theta_1 + 2(b_d + b_p) - a_1 + s$$

$$\iff \quad \theta_1 \geq b_d + s + 2b_p$$

(1)

That no jump occurs for any treatment above $a_2$ also implies that the doctor obtains his preferred treatment for all states greater than $\theta_2$.

We can now classify the equilibria in terms of $a_1$. If $a_1 > 2(b_d + b_p) + s$ then by Corollary 5 no equilibrium survives the intuitive criterion. If $a_1 \in [b_d, 2(b_d + b_p) + s]$ then $\theta_1 = a_1 + s + 2b_p$ satisfies condition (1) and we obtain equilibrium characterization (ii). If $a_1 \in (s, b_d]$ then $\theta_1$ is too small to satisfy condition (1), thus $\theta_1 = b_d + s + 2b_p$ is the smallest endpoint that can be used, and this obtains the characterization (i).
Equilibria in which treatment $s$ is induced

This equilibrium set is indexed by the highest state $\theta = \bar{s}$ for which treatment $s$ is induced. Once $\bar{s}$ is fixed, again the way an equilibrium may be constructed is substantially restricted.

**Lemma 6** The lowest state for which treatment $s$ is induced is $\underline{s} = \max\{0, s - b_d\}$.

**Proof** First suppose $s > b_d$. If $\underline{s} > s - b_d$ then there exists a type that prefers to induce $s$ and is not doing so, which is a contradiction. If $\underline{s} < s - b_d$, then types $\theta < \underline{s}$ either pool or separate. They cannot separate since there is a type $\theta \in (\underline{s}, s - b_d)$ that would defect from sending message $s$. If they pool then beliefs that reject off-the-equilibrium-path messages $m < s$ in the neighborhood of $s$ fail the intuitive criterion. When such a message is received the highest permissible belief is the doctor type indifferent between $s$ and $m$; however, if the doctor is indifferent then the patient strictly prefers the lower treatment and so accepts $m$.

If $s < b_d$ then the doctor’s preferred treatment is above $s$ for every state, and given that $s$ is induced for any state $\theta$, by Lemma 2 it must be induced for all states from 0 to $\theta$. ■

**Lemma 7** In any equilibrium that satisfies the intuitive criterion, the doctor induces his preferred treatment for $\theta < \underline{s}$.

**Proof** First, it is an equilibrium for $m(\theta) = \theta + b_s$ for $\theta < s - b_s$ because it is a best response for the doctor if the patient accepts and a best response for the patient since he infers the true state and prefers a treatment even lower than what the doctor prescribes. It is also the case that any treatment $a \in [b_d, b_d + s]$ cannot be off the equilibrium path, since that would imply the treatment is rejected but no beliefs satisfying the intuitive criterion could support this. ■

**Lemma 8** The treatment $s$ is induced on a set of positive measure, that is there exists $\bar{s} > \underline{s}$ so that $s$ is induced on $[\underline{s}, \bar{s}]$.

**Proof** Suppose not. There are two cases. First, the separating region may continue locally for $\theta > \underline{s}$. But this is impossible as the message $s + \varepsilon$ allows the patient to deduce the state is $s + \varepsilon - b_d$ to which $s$ is the patient’s best response. The second
case is that a next inducible treatment $a_1$ is discretely higher than $s$. But this fails
as there would exist types $s + \epsilon_2$ that would defect to the inducible message $s$. ■

We have thus established that $\bar{s} = \max\{0, s - b_d\}$ in every equilibrium, and the
choice in equilibrium construction begins with the choice $\bar{s}$.

**Proposition 10** Let $[\underline{s}, \bar{s}]$ denote the set of illnesses for which treatment $s$ is
induced. In any equilibrium $\bar{s} = \max\{0, s - b_d\}$. Furthermore, any equilibrium in which
treatment $s$ is induced has one of the following two forms:

i. If $\bar{s} \in (s, s + b_d + 2b_p]$ then the doctor prescribes

$$m(\theta) = \begin{cases} 
\theta + b_d & \theta \in [0, \bar{s}] \\
s & \theta \in [\bar{s}, \bar{s}] \\
2\bar{s} + 2b_d - s & \theta \in (\bar{s}, \bar{s} + b_d + 2b_p) \\
s + 4(b_d + b_p) & \theta \in [\bar{s} + 2b_d + 2b_p, s + 3b_d + 4b_p] \\
\theta + b_d & \theta \in [s + 3b_d + 4b_p, 1]
\end{cases}$$

ii. If $\bar{s} \in [s + b_d + 2b_p, \frac{1 + s - b_d}{2}]$ then the doctor prescribes

$$m(\theta) = \begin{cases} 
\theta + b_d & \theta \in [0, \bar{s}] \\
s & \theta \in [\bar{s}, \bar{s}] \\
2\bar{s} + 2b_d - s & \theta \in (\bar{s}, 2\bar{s} + b_d - s) \\
\theta + b_d & \theta \in [2\bar{s} + b_d - s, 1]
\end{cases}$$

If $\bar{s} \in (s + b_p, s + b_d + 2b_p)$ then no equilibrium survives the intuitive criterion.

**Proof** The type (i) equilibrium occurs for smaller values of $\bar{s}$ and has three pooling
regions while the type (ii) equilibrium occurs for larger values of $\bar{s}$ and has two pooling
regions. When $\bar{s} = s + b_d + 2b_p$, which is on the border between the two equilibrium
regions, the equilibrium looks the same when evaluated by either the type (i) or type
(ii) characterization. This is because in the type (i) characterization the top pooling
region shrinks to nothing.

First we consider $\bar{s} \in (s, s + b_d + 2b_p]$. Since at $\theta = \bar{s}$ the doctor’s preferred
treatment is strictly larger than $s$, there is a jump to the next highest treatment which
by Lemma 3 is $a_1 = 2(\bar{s} + b_d) - s$. In order for treatment $a_1$ to be accepted, again by
Lemma 3 it must pool at least out to $\theta_1 = \bar{s} + 2b_d + 2b_p$. By inspection, the doctor’s
preferred treatment at $\theta_1$ is above $a_1$ and thus there is a jump to $a_2 = s + 4(b_d + b_p)$.
Figure 5: An equilibrium in which $s = \frac{2}{10}$ is induced; $\bar{\pi} = \frac{3}{10}$, $b_p = b_d = \frac{1}{10}$.

Figure 6: An equilibrium in which $s = \frac{3}{10}$ is induced; $\bar{\pi} = \frac{1}{2}$, $b_p = \frac{1}{20}$, $b_d = \frac{1}{10}$.
The treatment \( a_2 \) is computed again using Lemma 3, and surprisingly is not a function of \( \bar{s} \). This observation is used in our later discussion of informativeness.

The pooling region at \( a_2 \) must extend out to at least the state for which \( a_2 \) is the doctor’s preferred treatment, namely \( \theta_2 = s + 3b_d + 4b_p \). By inspection, this makes the patient’s posterior strictly greater than what is necessary to induce him to accept. For states \( \theta > \theta_2 \), the doctor’s preferred treatments \( a(\theta) = \theta + b_d \) are induced.

The previous construction is well-defined up to \( \bar{s} = s + b_d + 2b_p \), at which point \( \theta_2 - \theta_1 = 0 \). For \( \bar{s} > s + b_d + 2b_p \), only two pooling regions can be supported and we find ourselves in a type (ii) equilibrium. Treatment \( a_1 = 2\bar{s} + 2b_d - s \) is constructed in the same manner as before. Previously, in order to induce acceptance it was necessary for the pooling region to extend beyond \( (a_d^*)^{-1}(a_1) \), thus creating another jump. However, now at \( (a_d^*)^{-1}(a_1) \) the patient’s posterior is larger than what is necessary for acceptance, hence the pooling region ends at \( \theta_1 = 2\bar{s} + b_d - s \), beyond which the doctor induces his preferred treatments for every state.

Lastly, when \( \bar{s} \) is sufficiently large the separating region disappears. This happens at \( \bar{s} = \frac{1}{2}(1 + s - b_d) \). Equilibria in this range are comprised of two pooling regions, one at \( s \) and one at a much higher treatment. It can be shown that these are easily ruled out on the basis of informativeness and we omit their characterization here. ■

**Corollary 6** An equilibrium in Proposition 10 survives the intuitive criterion unless \( s \in (s + b_p, s + b_d + 2b_p) \).

**Proof** By Corollary 5, no jumps may be supported above \( a = 2(b_d + b_p) + s \). This rules out type (i) equilibria in which \( a_1 > 2(b_d + b_p) + s \) but leaves in tact all type (ii) equilibria. ■

**Equilibrium Selection Using Informativeness**

Recall that previously we characterized equilibria using the smallest allowable pooling regions. For instance, in the class of equilibria in which treatment \( s \) is not induced, the smallest treatment \( a_1 \) was induced out to \( \theta_1 \) so that the patient’s posterior is sufficiently high to accept and the and doctor’s preferred treatment at \( \theta_1 \) was at least as large as \( a_1 \). Specifically, at least one of these constraints was required to bind. However, an equilibrium could also be supported in which this pooling region was extended and similar such decisions could have been made in other equilibrium classes.
It seems unlikely that increasing pooling beyond the constraints improves informativeness, and we demonstrate that doing so results in a mean-preserving spread in treatment. The following lemmas will apply to both classes of equilibrium discussed above.

**Lemma 9** Let \( a_1 \) be a pooled treatment ending at \( \theta_1 \) so that \( a_d^{*}(\theta_1) > 1 \). Let there also be only one pooling treatment \( a_2 > a_1 \) ending at \( \theta_2 \) at which \( a_d^{*}(\theta_2) = a_2 \). Then constructing an equilibrium with a higher \( a_1 \) creates a mean preserving spread in \( a(\theta) \). Similarly, if any pooled treatment \( a' \) can stop at \( (a_d^{*})^{-1}(a') \) but instead is extended beyond it such that another pooled treatment \( a'' > a' \) exists, then increasing the \( \theta' \), the endpoint of \( a' \), is a mean preserving spread.

**Proof** We first show \( E[\theta - a] \) is constant. Let \( \varepsilon > 0 \) be given and compare any \( \theta_1 \) to \( \theta_1 + \varepsilon \). It suffices to show the unconditional expected treatment on \( [\theta_1, 2\theta_1 + 2\varepsilon + b_d - a_1] \) is constant. This follows since

\[
(\theta_1 + b_d - a_1)(2\theta_1 + 2b_d - a_1) + \int_{2\theta_1 + b_d - a_1}^{2\theta_1 + b_d - a_1 + 2\varepsilon} (\theta + b_d) \, d\theta = \varepsilon a_1 + (\theta_1 + \varepsilon + b_d - a_1)(2\theta_1 + 2\varepsilon + 2b_d - a_1)
\]

We now show \( Var[\theta - a] \) increases. By construction, \( a_1 \) and \( a_2 \) are equidistant from \( \theta_1 + b_d \) and thus \( E[(\theta + b_d - a)^2] \) is higher for larger \( \theta_1 \) while \( E[(\theta + b_d - a)^2] \) is invariant. The results follows since \( Var[\theta - a] = Var[\theta + b_d - a] \). By a similar calculation, extending a pooled treatment \( a' \) beyond \( (a_d^{*})^{-1}(a') \) is also a mean preserving spread.

The previous lemma applies to equilibria in Proposition 9 for the selection of \( \theta_1 \) and \( \theta_2 \), to type (i) equilibria in Proposition 10 for the selection of \( \theta_1 \) and \( \theta_2 \), and to type (ii) equilibria in Proposition 10 for the selection of \( \theta_1 \).

**Proposition 11** Suppose \( b_d \leq s \). Then the equilibrium in Proposition 10 in which \( \bar{s} = s + b_p \) is the most informative equilibrium that survives the intuitive criterion.

**Proof** Recall that when \( b_d \leq s \) an equilibrium in which \( s \) is not induced cannot be supported. Thus we only examine equilibria in which \( s \) is induced. First suppose the equilibrium is of type (i), so that \( \bar{s} \in (\bar{s}, s + b_d + 2b_p) \). An explicit calculation reveals that the unconditional mean of \( a(\theta) \) is invariant to \( \bar{s} \), while the variance of \( \theta - a \) is
minimized at $\bar{s} = s + b_p$. For later comparison, we note here that $\text{Var}[\theta - a] = \frac{4}{3}(b_d + b_p)^3$. Now consider an equilibrium of type (ii), in which $s + b_d + 2b_p \leq \bar{s} \leq \frac{1}{2}(1 + s - b_d)$. By Lemma 9, increasing $\bar{s}$ in this region corresponds to a mean preserving spread. ■

**Corollary 7** The equilibrium in Proposition 8 of Krishna and Morgan (2001) (in which $b_p = 0$ and $\bar{s} = s$) is most informative.

**Proposition 12** Let $b_d > s$. Then, restricting attention to equilibria that survive the intuitive criterion, either the equilibrium in Proposition 9 in which $a_1 = b_d$ is most informative or the equilibrium in Proposition 10 in which $\bar{s} = s + b_p$ is most informative.

**Proof** In this parameter range both equilibria in which treatment $s$ is induced and equilibria in which treatment $s$ is not induced may be supported. Our strategy is to identify the most informative equilibrium, i.e. the one with the lowest $\text{Var}(\theta - a(\theta))$, in both classes and then compare the informativeness of the two winning equilibria. The results come from an explicit computation but we provide some intuition below.

First, for equilibria in which $s$ is not induced, as in Proposition 9, the most informative equilibrium is the one in which $a_1 = b_d$. Here, a higher value for $a_1$ increases the interval $[0, \theta_1]$ and keeps the interval $[\theta_1, \theta_2]$ constant, which ends up resulting in higher variance. For a lower value of $a_1$, the interval $[0, \theta_1]$ remains the same and the interval $[\theta_1, \theta_2]$ increases. An explicit computation confirms that $a_1 = b_d$ is the most informative and yields

$$\text{Var}(\theta - a(\theta)) = \frac{2}{3}(b_d + 2b_p + s)^3$$

(2)

Next, for equilibria in which $s$ is induced, as in Proposition 10, the most informative equilibrium is the one in which $\bar{s} = s + b_p$. Here, it can be seen using Lemma 9 that any move above or below $\bar{s} = s + b_p$ is a mean preserving spread. Again, an explicit computation reveals that

$$\text{Var}(\theta - a(\theta))$$

$$= -\frac{1}{4}s^4 + s^3b_d + \frac{1}{3}s^3 - \frac{3}{2}s^2b_d^2 - s^2b_d + sb_d^3 + sb_d^2 - \frac{1}{4}b_d + b_d^3 + 4b_d^2b_p + 4b_dp^3 + \frac{4}{3}b_p^3$$

(3)
Lastly, comparing these two expressions we find that either can dominate the other depending on bias parameters $b_d$ and $b_p$. ■

**Corollary 8** If $b_p > b_d > s$ then the most informative equilibrium that survives the intuitive criterion is the equilibrium from Proposition 9 in which $a_1 = b_d$.

The proof follows from comparing expressions 2 to 3 when $b_d = b_p$ and then taking a derivative with respect to $b_p$.

**Appendix B: Other Proofs**

**Proof of Corollary 4** We demonstrate the existence of $p$ and $c$ such that $U_p(a_p^*) + U_p(a_{eqm}) < 2U_p(a_d^*)$. Let $p = c = \frac{1}{4}$ so that $3p - c = \frac{1}{2}$. A calculation shows

$$2U_p(a_d^*) = \frac{1}{4} > \frac{25}{96} = -\frac{3}{32} - \frac{1}{6} = U_p(a_p^*) + U_p(a_{eqm}).$$

Because the payoff functions are continuous, the parameter space for which the inequality holds has positive measure. ■

**Proof of Proposition 4** In our model the difference in actions grows with the interval of over- and undertreatment and expected utility declines in it. Thus it suffices to track increases in the interval on which over- and undertreatment occur. Expected undertreatment is $(3p - c) - (2p - c) = p$, which increases in $p$, is constant with respect to $c$, and increases for an equal increase in $p$ and $c$ (fixed markup). Expected overtreatment is $(6p - 2c) - (3p - c) = 3p - c$, which increases in $p$, decreases in $c$, and increases for an equal increase in $p$ and $c$ (fixed markup).

$$U_d(a(\theta)) = -\frac{1}{2}E[(a(\theta) - \theta)^2] + (p - c)E[a(\theta)]$$

$$U_p(a(\theta)) = -\frac{1}{2}E[(a(\theta) - \theta)^2] - pE[a(\theta)]$$

These functions share the quadratic loss term and so their derivatives will share the terms $\frac{d}{dc}E[(a(\theta) - \theta)^2] = -2(3p - c)^2$ and $\frac{dp}{dE}[a(\theta) - \theta)^2] = 6(3p - c)^2$. A
calculation shows

\[
\begin{align*}
\frac{dU_p}{dc} &= (3p - c)^2 + (p - c) + p > 0 \\
\frac{dU_p}{dp} &= -3(3p - c)^2 - p - \frac{1}{2} < 0 \\
\frac{dU_d}{dc} &= (3p - c)^2 - (p - c) - \frac{1}{2} < 0 \\
\frac{dU_d}{dp} &= -3(3p - c)^2 + (p - c) + \frac{1}{2} \geq 0
\end{align*}
\]

where this last term is ambiguous since, for example, letting \( p = c \) it is positive when \( p < \frac{1}{2\sqrt{6}} \) but negative when \( p > \frac{1}{2\sqrt{6}} \). It follows from above that \( \frac{dU_d}{dc} + \frac{dU_d}{dp} < 0 \) and \( \frac{dU_p}{dc} + \frac{dU_p}{dp} < 0 \).

**Proof of Proposition 6** First note that total expected treatment is constant in \( \gamma \),

\[
\begin{align*}
&= \int_0^{(1+2\gamma)p-c} (p - c) d\theta + \int_{(1+2\gamma)p-c}^{(2+4\gamma)p-2c} ((3 + 4\gamma)p - 3c) d\theta + \int_{(2+4\gamma)p-2c}^1 (\theta + p - c) d\theta \\
&= p - c + \frac{1}{2}
\end{align*}
\]

and thus \( \frac{\partial E[a]}{\partial \gamma} = 0 \). It therefore suffices to show that the treatment is more accurate, i.e., that \( \int_0^1 - (\theta - a)^2 d\theta \) is decreasing in \( \gamma \). This follows from differentiating

\[
\begin{align*}
&\int_0^{(1+2\gamma)p-c} -\frac{1}{2} (\theta - (p - c))^2 d\theta + \int_{(1+2\gamma)p-c}^{(2+4\gamma)p-2c} -\frac{1}{2} (\theta - ((3 + 4\gamma)p - 3c))^2 d\theta \\
&\quad + \int_{(2+4\gamma)p-2c}^1 -\frac{1}{2} (\theta - (\theta + p - c))^2 d\theta
\end{align*}
\]

with respect to \( \gamma \) to find \( -2p(p - c + 2p\gamma)^2 < 0 \) since \( p \geq c \). Thus \( \gamma = 0 \) is optimal for the patient.

**Proof of Proposition 7** Type \( \theta_1 = 3p - c \) has most incentive to defect from any equilibrium as his equilibrium action is farthest from his preferred. It suffices to find \( \alpha \) such that this type is indifferent between his equilibrium payoff and the payoff obtained from defecting and having \( \alpha \) consumers take action \( a^*_d(\theta_1) \) and the

34
remaining \((1 - \alpha)\) take action zero:

\[
u_d(\theta_1, a = p - c) = \alpha u_d(\theta_1, a = \theta + p - c) + (1 - \alpha) u_d(\theta_1, a = 0)\]

A solution \(\alpha^* \in (0, 1)\) exists as \(U_d(\theta_1, a = 0) < U_d(\theta_1, a = p - c) < u_d(\theta_1, a = \theta + p - c)\). Thus whenever \(\alpha \leq \alpha^*\) our equilibrium survives unchanged. ■

References


