ISSN 2281-4299

DIPARTIMENTO DI INGEGNERIA INFORMATICA
AUTOMATICA E GESTIONALE ANTONIO RUBERTI

SAPIENZA
UNIVERSITÀ DI ROMA

Privacy protection and AI-based pharmaceutical innovation: Friends or foes?

Chiara Conti
Anna D'Annunzio
Pierfrancesco Reverberi

Technical Report n. 01, 2025

Privacy protection and AI-based pharmaceutical innovation: Friends or foes?

Chiara Conti,[a] Anna D'Annunzio,[b] Pierfrancesco Reverberi[a,c]

Abstract

We study private incentives to invest in a targeted treatment for an eligible patient group based on collected health data, and compare them with the social optimum. Patients must be compensated for sharing personal data, because they incur an idiosyncratic privacy cost and can partially free ride on data-driven health innovation. We find that, when (for a given compensation) data collection costs are higher (lower) than aggregate privacy costs, the firm underinvests (may overinvest) in the new treatment. Thus, in equilibrium, underinvestment arises with uniform compensation to patients, whereas overinvestment may arise with personalized compensations. We also find that privacy protection measures can align investment incentives with social goals. Then, we consider an alternative scenario where a public agency may provide the firm respectively with: (i) free access to all patients' health data; (ii) access to health data conditional on a lump-sum payment; (iii) free access to a selected amount of health data. Centralized governance of health data removes externalities in personal data sharing. We generally find that this can foster investment in the new treatment and improve social welfare.

Keywords: Health data; Data-driven innovation; Precision medicine; Privacy regulation

[a] Department of Computer, Control, and Management Engineering, Sapienza University of Rome.

[b] Tor Vergata University of Rome.

[c] Corresponding author: Via Ariosto 25, 00185 Roma; +39 06 77274096; reverberi@diag.uniroma1.it.

1. Introduction

The potential benefits of artificial intelligence (AI) techniques in the field of healthcare are widely recognized as extremely important.[1] Indeed, AI models and algorithms support faster drug innovation as well as the design of targeted treatments based on patients' genetic, environmental, and lifestyle features (also called *precision medicine*). By customizing therapies to the unique characteristics of each patient, targeted treatments may minimize adverse effects and enhance survival rates in critical health conditions, such as cancer, compared to conventional treatments that rely on a standardized approach designed for the "average patient" (World Health Organization, 2021).

The efficacy of AI systems for pharmaceutical R&D strongly hinges on the availability of a vast amount of medical data. Nonetheless, strict rules often govern the collection and use of patients' health data to develop innovative treatments. In Europe, the General Data Protection Regulation (GDPR) has provided individuals with control over personal data. The GDPR imposes severe requirements on the handling of health data due to its classification as "special category data". These requirements are intended to protect individual privacy, avoid risk of discrimination, ensure trust in healthcare systems, and prevent potential misuses of sensitive data.

At the same time, the European Commission (EC) has deemed that lighter requirements to gather and employ health data can improve pharmaceutical R&D, which may be hindered if it is too costly to obtain consent from each patient to use personal data for medical research. Hence, to encourage health innovation, the EC has recently proposed to create a supra-national health data governance system, namely, the European Health Data Space (EHDS), where patients' data can be accessed after permission of a public agency (EC, 2022).

In this paper, we study the interplay between data-driven pharmaceutical innovation and privacy regulation. We address the following questions. What is the impact of privacy regulation on medical innovation? More precisely, does a consent-based regime for patients' health data collection and use work against the development of targeted treatments? Does enhanced privacy protection ultimately benefit patients and society

[1] For instance, roughly 800,000 Americans each year have been estimated to suffer from poor medical decision making. AI systems are expected to enhance diagnostic accuracy and disease tracking, improve prediction of patients' outcomes and suggest more effective treatments (see The Economist, *The AI doctor will see you…eventually*, March 27, 2024, https://www.economist.com/leaders/2024/03/27/the-ai-doctor-will-see-youeventually, and MIT Technology Review, 2018. *The Precision Medicine Issue*, https://www.technologyreview.com/magazines/the-precision-medicine-issue/).

as a whole (when taking into account the effects on incentives to invest in targeted treatments)? Which forms of centralized government of patients' health data, if any, can foster investment in new treatments and improve social welfare?

We build a theoretical model to analyze incentives of a pharmaceutical firm to develop a targeted treatment for an eligible patient group, rather than offer a standard treatment for all patients.[2] The new treatment may be achieved by investing in a stochastic R&D process whose likelihood of success rises with the total amount of health data collected by the firm. Because patients are privacy sensitive, the firm must compensate each of them to obtain consent to collect personal data for medical research.

We assume that patients in the eligible group may benefit from the targeted treatment (if achieved) even if they do not share health data. This creates a free-riding issue due to the public good nature of the new treatment. Indeed, eligible patients wish to receive the treatment without incurring privacy costs for sharing data. We also assume that patients are more likely to receive the new treatment when they share data than when they refuse to share. The reason is that genetic profiles are identified at the time when research is performed for patients sharing data, whereas only at a later time for patients not sharing data, possibly through biomarker tests with limited precision or accessibility.

It is worth noting that the probability to identify eligible patients who do not share data creates a trade-off in developing the new treatment. On the one hand, when this probability is low patients have an incentive to share data even for a limited compensation. On the other hand, for a given amount of collected data, a lower probability to identify eligible patients among those not sharing data reduces the market size for the targeted treatment, thereby negatively affecting revenues from the treatment.

We assume that drug prices are negotiated with the government, which fully reimburses patients.[3] We compare equilibrium incentives to invest in the targeted treatment with the social optimum, depending on four main factors: i) incremental health benefit of the targeted treatment; ii) eligible patient group size; iii) privacy costs for patients; iv) degree of free riding in sharing health data. These factors, in turn, affect the value of

[2] Gonzalez *et al*. (2016) define (radical) "horizontal" drug innovations as advances benefiting a given patient group because of lower side effects. They provide several examples of such innovations in the market for statins.

[3] This assumption fits oncological, genetic, and degenerative diseases in national health systems of many EU countries.

patients' compensation for data sharing, and thus the total cost of data collection for the firm, as well as the aggregate privacy costs for patients who share health data.

In the baseline model, we assume that the firm knows the distribution of privacy costs across patients but does not know the idiosyncratic privacy cost of each patient. Thus, she offers a uniform compensation to induce patients to share health data. We find that the firm underinvests in the targeted treatment compared to the social optimum. More specifically, either the firm does not invest in data collection, thus in the new treatment, when it would improve welfare, or she collects a smaller amount of health data than socially optimal, thereby reducing the likelihood to obtain the new treatment.

We argue that the rationale for this result is that, for a given compensation to patients, total costs of health data collection are higher than aggregate privacy costs. It follows that, *ceteris paribus*, lower data collection costs may align private and social incentives to invest in the new treatment. An intuitive way to achieve this goal is by adopting privacy protection measures that reduce patients' perceived costs of sharing health data.

We study whether and how this result depends on the information available to the firm. For this purpose, we consider the opposite case to the baseline model where the firm knows the privacy cost of each patient (e.g., because it can be inferred from individuals' observable attributes such as general health conditions or socio-economic status). In such a case, the firm can offer personalized compensations to patients thereby saving on data collection costs, which (for a given compensation to the marginal patient sharing health data) become smaller than aggregate privacy costs. We thus find that, compared to the social optimum, the firm may overinvest in the targeted treatment if she has strong bargaining power in price negotiation with the government. The rationale is that the firm can reap most of the benefits from developing the targeted treatment.

Then, we consider an alternative scenario where a central (public) agency controls patients' health data and the use of such data for R&D purposes, while patients are not compensated for sharing data. Centralized governance of health data removes externalities in personal data sharing, which determine underinvestment with uniform compensation. In this framework, we investigate the effects of several policy measures meant to align private and social incentives to invest in the new treatment.

First, we assume that the central agency provides the firm with free access to patients' health data. We find that this policy may yield overinvestment in the targeted treatment when the firm's bargaining power is high enough. Indeed, the firm finds it profitable to invest using all patients' data whereas, from a welfare

perspective, it would be optimal to invest using less data, or not to invest at all. Intuitively, overinvestment more likely occurs when privacy costs are high. We also find that social welfare improves with free access to all patients' data relative to the baseline model when the total incremental health benefit of the targeted treatment is sufficiently high. In addition, free access to health data is more likely to improve welfare when patients' privacy costs are low.

Second, we assume that the firm is required to make a lump-sum payment to the central agency to gain access to all patients' health data. Compared to the case where the firm has free access to health data, this policy reduces overinvestment in the targeted treatment. Moreover, it makes more likely that social welfare is improved relative to the baseline model.

Third, we consider a policy granting free access to an amount of health data selected by the agency. We show that this policy achieves the first-best outcome when the firm has enough bargaining power, or when the total incremental health benefit of the new treatment is sufficiently high. If such benefit takes intermediate values, then the agency decides to grant free access to a larger amount of health data than socially optimal to induce the firm to invest in the targeted treatment. As a result, social welfare improves relative to the baseline model.

This paper is organized as follows. Section 2 discusses the literature. Section 3 presents the model. Section 4 derives the equilibrium. Section 5 analyzes welfare, while Section 6 compares private and social incentives to invest in targeted treatments. Section 7 considers the case with personalized compensations. Section 8 discusses policy measures, and Section 9 concludes.

2. Literature review

Our paper connects two related strands of literature. The first one considers the interaction between different forms of drug price regulation and investment incentives in pharmaceutical R&D (Bardey *et al*., 2010; Gonzalez *et al*., 2016; Brekke *et al*., 2023). Specifically, this literature investigates how regulation affects the incentives of innovative firms to conduct R&D activities, the intensity of R&D and the types of innovations that are brought to the market.

The second strand studies the relationship between privacy regulation and product innovation (Conti and Reverberi, 2021; Lefouili *et al*., 2024). It shows that the effects of a consent-based privacy regime on the

quality of products and on consumer surplus depend on the strength of the complementarity between personal data sharing and product quality.

Rhodes and Zhou (2024) introduce a general framework of personalization and privacy choices and apply it to (among others) personalized product design. They focus on negative externalities imposed by consumers who share personal data on those who do not share. Indeed, they find that firms degrade their product offering to non-sharing consumers to extract more surplus from sharing ones through personalized products sold at higher prices. Overall, too many consumers share data.

Our model includes distinctive features of healthcare markets, such as price negotiation between the firm and the government, and considers patients' compensation for sharing health data. Hence, in contrast to Rhodes and Zhou (2024), we find that too much data sharing (and thus overinvestment in the new treatment) may occur even with positive externalities across individuals.

A very recent paper that analyzes personal data sharing issues in healthcare is Canta *et al*. (2025). They consider a two-sided platform that offers health services to patients and may collect their data to share it with innovators for a price. Specifically, they assess the effects of different privacy policies on the incentives to sell data and invest.

Despite the widespread interest in precision medicine (see, e.g., Stern *et al*., 2017), the theoretical literature on this topic is limited. Antoñanzas *et al*. (2015) study the incentives of health authorities to use predictive biomarkers to inform treatment choices. Brekke *et al*. (2024) examine how biomarker tests affect both competition between existing drugs and the design of health plans. Mougeot and Naegelen (2022) assess the impact of price regulation on the viability of precision medicine. Nonetheless, all these papers ignore privacy issues related to personal data sharing, which are key to our analysis. On an empirical ground, Miller and Tucker (2018) find that patients' control over data redisclosure, but not an informed consent policy, boosts the spread of genetic testing.

The paper closest to ours is Conti *et al*. (2024), which investigates private investment in data-driven drug innovation in the limit case of the highest degree of free riding (where eligible patients receive full benefits of the targeted treatment even if they do not share personal data). In this paper, we generalize the baseline model by parameterizing the degree of free riding. Furthermore, we introduce the case with personalized (instead of

uniform) compensations. Finally, we explore the rationale and effects of different forms of centralized governance of health data.

3. The model

We consider a therapeutic market where a pharmaceutical firm offers a "one-size-fits-all" treatment to patients. The firm can invest in R&D to develop a targeted treatment with a higher health benefit for a group of patients. For this purpose, the R&D process requires as input patients' health data. Such data are not freely available to the firm because privacy regulation protects patients by providing them with control over personal data. Thus, the firm may have to compensate privacy-sensitive patients to obtain consent to collect personal health data for research purposes.

We develop a game-theoretic model to study the interaction among patients, the pharmaceutical firm, and the government.

Patients

There is a unit mass of heterogeneous patients in the therapeutic market. Patients differ across two orthogonal dimensions: genetic features and privacy attitudes.

As to genetic features, we distinguish two groups of patients. The first is a homogeneous group which includes patients who are eligible for the targeted treatment. We assume that the size of this group is α, with $0 < \alpha < 1$. Then, the complementary group has size $(1 - \alpha)$ and includes the rest of (possibly heterogeneous) patients in the therapeutic market who are not eligible for the targeted treatment. We assume that the size α of the eligible group is common knowledge.[4] Thus, patients know the probability of belonging to this group as they may be informed by the pharmaceutical firm or the physician. However, individual patients do not know *ex ante* (i.e., before the firm develops the targeted treatment) whether they are part of the eligible group.

As to privacy attitudes, there is a continuum of patients' types with idiosyncratic privacy costs. Each privacy type is identified by a privacy cost $\beta > 0$, which reflects a taste for privacy. This is likely to be correlated with

[4] According to data-driven pharmaceutical R&D, the size of the eligible group of patients (to be targeted with a specific treatment) can be estimated using past empirical observations, such as responses to standard therapies, and investigating networks of available scientific and patients' data by means of AI techniques.

individual characteristics, health conditions, and socio-economic status. We assume that privacy types are uniformly distributed over the interval $[0,\ \bar{\beta}]$. Patients are aware of their privacy types.

Each patient has an average health benefit $q_0 > 0$ from the standard treatment. If the firm invests in R&D and the investment is successful, the targeted treatment can provide eligible patients with a health benefit q_H, with $q_H > q_0$ (e.g., due to lower side effects or higher probability of recovery). We assume that, if a patient decides to share health data, then the probability of receiving the targeted treatment is α (i.e., the eligible group size). Instead, if the patient does not share health data, then she has a lower probability to receive the targeted treatment, which is measured as $\varepsilon\alpha$, with $0 < \varepsilon \leq 1$.

Parameter ε can be related, for instance, to the precision of the biomarker test that is developed jointly with the targeted treatment to identify eligible patients among those who are not sharing health data. Alternatively, it can be interpreted as a discount factor, because eligible patients who do not share data might be able to receive the targeted treatment later than eligible patients who share data.

Given perceived privacy costs, the firm offers patients a compensation δ for sharing health data. This may occur through monetary payments or health benefits from participation in medical trials. Thus, the utility for a patient with privacy type β depends on two components: the health benefit for the received treatment (i.e., the first term in squared brackets in equation (1) below) and, conditional on data sharing, the privacy surplus, that is the difference between the compensation δ for sharing data and the privacy cost (i.e., the second term in squared brackets in (1), which is equal to zero if the patient does not share data).

Let ϕ be the probability that investment in the targeted treatment is successful. Thus, the expected utility for a patient with privacy type β is:

$$E(U(\delta,\beta)) = \begin{cases} [\phi(\alpha q_H + (1-\alpha)q_0) + (1-\phi)q_0] + [\delta - \beta] & if\ \beta\ shares\ data \\ [\phi(\alpha\varepsilon q_H + (1-\alpha\varepsilon)q_0) + (1-\phi)q_0] & if\ \beta\ does\ not\ share \end{cases} \quad (1)$$

Monopolist

A monopolist offers a "one-size-fits-all" treatment with an average health benefit of q_0 to all patients in the therapeutic market. For simplicity, we assume production costs are zero. The firm may invest in project H, which may yield a targeted treatment to the eligible patient group of size α. For this purpose, the firm needs to collect patients' health data and combine these data with other medical and non-medical information to

identify the specific genetic features of the group (for instance, to find out genetic mutations responsible for heterogeneity in responses to standard therapies).[5]

We assume that project H is stochastic, and the probability of success $\phi(d)$: $[0,1] \to [0,1]$ increases with the amount d of patients' health data (i.e., $\phi'(d) > 0$, with $\phi(0) = 0$ and $\phi(1) = 1$). For simplicity, we assume that $\phi(d) = d$. In turn, the amount of collected data depends on the compensation offered to patients.

In the baseline model, we assume that patients' privacy types β are private information, and the firm only knows the distribution of values of β across patients. Thus, the firm chooses a uniform compensation δ to be offered to each patient to collect the desired amount of health data.[6] Hence, the cost of data collection entailed by project H is $c(\delta) = \delta d$. We assume that the total cost for undertaking project H is $F + c(\delta)$, where $F > 0$ is the R&D cost.

It follows that the expected profit from investment in project H can be written as:

$$\mathrm{E}(\Pi_H) = \phi\big(d(\alpha p_H + (1-\alpha)p_0) + (1-d)(\varepsilon\alpha p_H + (1-\varepsilon\alpha)p_0)\big) + (1-\phi)p_0 - c(\delta) - F \quad (2)$$

where p_H is the price of the targeted treatment and p_0 is the price of the "one-size-fits-all" treatment.

Instead, if the monopolist does not invest in project H and thereby offers the standard treatment to all patients, then profit is simply $\Pi_0 = p_0$.

Government

Consistent with the relevant literature (see, e.g., Bardey *et al*., 2010), we assume that, conditional on successful innovation, the government negotiates the price of the targeted treatment p_H with the firm through Nash bargaining. On the other hand, we consider the price of the standard treatment p_0 to be exogenous (e.g., because it has been negotiated in the past). We also assume that treatments are entirely paid by the government (see footnote 3). Hence, the therapeutic market is fully covered.

Timing

The timing of the game is as follows:

1. The monopolist decides whether to undertake project H. Under project H, the firm chooses the compensation for patients, and thereby how much health data to collect.

[5] See, e.g., the case studies reported in Gavan *et al*. (2018).

[6] In Section 7, we consider the case where the firm has complete information on patients' privacy types and thereby offers personalized compensations.

2. Given that the firm has undertaken project H and set the compensation for patients, each patient decides whether to share personal health data.
3. If the targeted treatment is achieved, then the firm and the government negotiate the price of the treatment through a (generalized) Nash bargaining process.[7]

We solve the game backwards for subgame perfect Nash equilibria. We use superscript * to denote equilibrium variables and superscript *W* to denote the social optimum. Proofs are in the Appendix.

4. Equilibrium

In this section, we derive the equilibrium of the game, and we study the incentives to invest in project H. We show that these incentives depend essentially on the importance of privacy costs and on the maximum total incremental health benefit $\alpha(q_H - q_0)$ of the targeted treatment, which, in turn, is obtained from the individual incremental benefit $(q_H - q_0)$ for the treatment and the size α of the eligible group of patients. In the next section, we derive the social optimum, which in Section 6 we compare with the equilibrium of the game.

In what follows, we make the following assumptions.

Assumption 1. *Let* $\alpha(q_H - q_0) < \frac{\bar{\beta}}{2(1-\varepsilon)}$.

Assumption 2. *Let* $\bar{\beta} > \max\left\{F\left(1 + \frac{2(1-\varepsilon)(2-\lambda)}{\varepsilon(1-\lambda)}\right), \frac{2F(2-\varepsilon)}{\varepsilon}\right\}$.

Assumption 1 ensures concavity of both welfare and profit functions.[8] Assumption 2 ensures that an interior solution may exist, thereby excluding that there are only corner solutions in the welfare and profit maximization problems. Indeed, if privacy costs are low enough (such that Assumption 2 does not hold) then, depending on the fixed R&D cost and on the bargaining power in price negotiation, the firm either uses all patients' health data to invest in project H or does not invest at all. The same holds for the government.

Stage 3. Price negotiation

[7] We assume that the price of the targeted treatment is negotiated *ex post*, namely, after the treatment is achieved, given that the firm has invested in R&D and collected health data (so that R&D as well as data collection costs are sunk). This timing is widely accepted in the relevant literature (see, e.g., Bardey *et al*., 2010; Barigozzi and Jelovac, 2020; Matteucci and Reverberi, 2017). This is because a pharmaceutical company obtains patent protection and the legal authorization to manufacture and sell the drug only after the innovation occurs.

[8] Assumption 1 also rules out the case where all patients are willing to share data without compensation (i.e., for $\delta = 0$).

We assume that in the *status quo* all patients receive the standard treatment, which is sold at price p_0, with a health benefit of q_0. Price negotiation for treatment H occurs through a Nash bargaining process where, in the absence of an agreement, eligible patients do not receive the new treatment.

Given that R&D and data collection costs as well as patients' privacy surplus are sunk at the price negotiation stage, we assume that the government cares about *incremental* patients' benefits relative to the *status quo* net of social expenditure. More formally, the government considers $\Delta\widetilde{W} = \alpha[(q_H - q_0) - (p_H - p_0)][d + (1-d)\varepsilon]$. On the other hand, the firm cares about the *incremental* profit relative to the *status quo*, namely, $\Delta\Pi = \alpha(p_H - p_0)\,[d + (1-d)\varepsilon]$.

Let $\lambda > 0$ (respectively, $(1-\lambda) > 0$) be the government's (respectively, firm's) bargaining power in price negotiation. Then, the Nash bargaining problem is:

$$\max_{p_H} \left(\Delta\widetilde{W}\right)^{\lambda} (\Delta\Pi)^{1-\lambda},$$

which yields the negotiated price for treatment H:[9]

$$p_H = p_0 + (1-\lambda)(q_H - q_0). \tag{3}$$

Stage 2. Data sharing

We assume that patients do not know *ex ante* whether they are eligible for the targeted treatment. However, they are informed on the probability of receiving the targeted treatment (i.e., α if they decide to share health data, and $\varepsilon\alpha$ if they do not share). Furthermore, they are aware that the probability of successful innovation in the targeted treatment increases proportionally with the total amount of collected health data.

Given that the firm undertakes project H, a patient with privacy type β accepts compensation δ and shares personal health data if the expected utility is at least as high as in the alternative case where she does not share data. Let ϕ^e be the expected probability of success of project H. Formally, the patient with privacy type β decides to share health data if and only if:

$$\phi^e\{\alpha q_H + (1-\alpha)q_0\} + (1-\phi^e)\{q_0\} + \delta - \beta \geq \phi^e\{\alpha\varepsilon q_H + (1-\alpha\varepsilon)q_0\} + (1-\phi^e)\{q_0\}.$$

The above inequality simplifies to:

$$\beta \leq \phi^e\,\{\alpha(1-\varepsilon)(q_H - q_0)\} + \delta. \tag{4}$$

[9] We find that the negotiated price p_H is independent of α and ε. The rationale is that, at this stage of the game, the size of the market for the targeted treatment, namely, $\alpha[d + (1-d)\varepsilon]$ is fixed.

Rearranging (4), we obtain $\delta \geq \beta - \phi^e \{\alpha(1-\varepsilon)(q_H - q_0)\}$. Hence, the minimum compensation δ that must be offered to collect health data from a given patient is lower than her privacy cost β, except for the limit case where $\varepsilon = 1$ (when such a compensation exactly reflects the privacy cost).

It follows from (4) that the patient with the highest privacy cost $\bar{\beta}$ shares health data as long as $\bar{\beta} \leq \alpha(1-\varepsilon)(q_H - q_0) + \delta$, since $\phi^e = 1$ when all patients share data. In turn, this implies that $\bar{\delta} = \bar{\beta} - \alpha(1-\varepsilon)(q_H - q_0)$ is the (maximum) compensation the firm is willing to offer to collect data from all patients.[10] Lower values of δ make patients with high privacy costs unwilling to share data.

Let $\hat{\beta}$ be the privacy type who is indifferent between sharing health data or not. In equilibrium, patients' expectations need to be fulfilled, that is, the expected probability of success of project H must be equal to the value of the probability function when the amount d of collected data is determined by the indifferent patient (i.e., the marginal patient who shares data). This implies that $\phi^e = d^e = \frac{\hat{\beta}}{\bar{\beta}}$. It follows from (4) that $\hat{\beta} = \frac{\hat{\beta}}{\bar{\beta}}\{\alpha(1-\varepsilon)(q_H - q_0)\} + \delta$. Hence, under fulfilled expectations, the indifferent privacy type is given by:

$$\hat{\beta} = \delta\bar{\beta}/\bar{\delta}, \tag{5}$$

and patients with privacy types β such that $\beta \leq \hat{\beta}$ decide to share health data.

From (5) we find that, for a given δ, $\hat{\beta}$ is decreasing in ε, meaning that a higher probability for patients to receive the new treatment without sharing personal data reduces the total amount of collected health data. In this set up, the targeted treatment has some features of a public good relative to health data, and this causes a free-riding issue, because individual patients would aim at receiving the benefits of the treatment without incurring the (privacy) costs of sharing data. In this sense, parameter ε can be interpreted as patients' degree of free riding on health data sharing.

Stage 1. Data collection and investment

In equilibrium, the firm invests in the targeted treatment if the expected profit under project H is at least as high as the profit from selling only the standard treatment, that is, $E(\Pi_H) \geq \Pi_0$. This requires the expected probability of success of project H to be high enough. In turn, this means that the firm should collect a sufficient amount of health data (by offering a compensation $\delta > 0$ to patients).

[10] A higher value of δ than the value for which all patients share health data is neither profitable nor socially optimal.

For a given value of δ (such that $\delta \leq \bar{\delta}$), from (5) the probability of success of project H is $\phi = d = \delta/\bar{\delta}$, and thereby the cost of data collection is $c(\delta) = \delta^2/\bar{\delta}$. Thus, plugging (3) into (2), the monopolist's problem can be written as:

$$\max_{\delta} E(\Pi_H) = \frac{\delta}{\bar{\delta}}\alpha(1-\lambda)(q_H - q_0)\left(\frac{\delta}{\bar{\delta}}(1-\varepsilon)+\varepsilon\right) - \frac{\delta^2}{\bar{\delta}} + p_0 - F, \tag{6}$$

where the monopolist collects health data if and only if $E(\Pi_H) \geq \Pi_0$, and an interior solution exists as long as $0 < \delta^* < \bar{\delta}$.

It is worth noting from (6) that parameter ε has contrasting effects on incentives to invest in the new treatment. On the one hand, (*ceteris paribus*) low values of ε limit the degree of free riding and encourage patients to share personal data. Hence, the firm can reduce patients' compensation to collect health data. On the other hand, low values of ε reduce the size of the market for the targeted treatment, because the firm hardly identifies patients who do not share data. Hence, the firm has the incentive to increase the compensation offered to patients to mitigate this negative effect.

Let $A \equiv \alpha(q_H - q_0)$ be the maximum total incremental health benefit of the targeted treatment. Proposition 1 determines the conditions for which the firm finds it profitable to offer a compensation to patients for collecting health data, thereby investing in project H.

***Proposition* 1.** *Let* $\underline{A^H} = \frac{-2F(1-\varepsilon)(2-\lambda)+2\sqrt{[F(1-\varepsilon)(2-\lambda)]^2+F\bar{\beta}\varepsilon^2(1-\lambda)^2}}{\varepsilon^2(1-\lambda)^2}$ *and* $\overline{A^H} = \frac{2\bar{\beta}}{[\varepsilon(1-\lambda)+2(1-\varepsilon)(2-\lambda)]}$. *We have that:*

(i) when $\underline{A^H} \leq A < \overline{A^H}$, *the firm offers* $\delta^* = \frac{\bar{\delta}\varepsilon(1-\lambda)A}{2\left(\bar{\beta}-A(2-\lambda)(1-\varepsilon)\right)}$ *to patients, and the amount of collected health data for project H is* $d^* = \frac{\varepsilon(1-\lambda)A}{2\left(\bar{\beta}-A(2-\lambda)(1-\varepsilon)\right)} < 1$;

(ii) when $A \geq \overline{A^H}$, *the firm offers* $\delta^* = \bar{\delta}$ *to patients, and collects all patients' health data for project H, so as* $d^* = 1$;

(iii) when $A < \underline{A^H}$, *investment in project H is not profitable (so as* $\delta^* = d^* = 0$*).*

Proposition 1 states that the firm is willing to collect data and invest in R&D if A is high enough. Indeed, for low values of the total incremental health benefit of the targeted treatment, that is, $A < \underline{A^H}$, the firm does not find it profitable to invest because $E(\Pi_H) < \Pi_0$. We find that the critical value $\underline{A^H}$ for investment in project H is increasing in $\bar{\beta}$. The higher the value of the upper bound for privacy costs $\bar{\beta}$, the higher the dispersion of

privacy concerns across patients and thus the average privacy concern. Hence, when patients are more privacy concerned, the targeted treatment is less likely to be developed. Indeed, health data are essential to R&D investment, but are costly to collect.

Instead, when the total incremental health benefit of the new treatment is very high, that is, $A \geq \overline{A^H}$, the firm finds it profitable to compensate patients so as to induce all of them to share health data. As $\bar{\beta}$ increases, the compensation $\bar{\delta}$ necessary to induce all patients to share data increases as well, thereby making it more difficult for the firm to collect all patients' data.

In the following section, we derive the optimal compensation to patients and incentives to invest in the targeted treatment from a welfare perspective, so as to evaluate the desirability of private investment in the new treatment from a social point of view.

5. Welfare

Social welfare under treatment H, W_H, depends on patients' total health benefit from the new treatment $\widehat{W}_H$, the fixed R&D cost F, and aggregate privacy costs PC. Indeed, the social expenditure and the firm's revenues for treatment i (with $i \in \{0, H\}$) cancel out in the welfare function. The same occurs for the cost of data collection for the firm and the total compensation for patients for data sharing under investment in project H (both of these components are zero without investment). Thus, the expected welfare under treatment H is $E(W_H) = E(\widehat{W}_H) - F - PC$. More specifically, we have:

$$E(\mathrm{W}_H) = \phi\left(d\big(\alpha(q_H) + (1-\alpha)(q_0)\big) + (1-d)\big(\varepsilon\alpha(q_H) + (1-\varepsilon\alpha)(q_0)\big)\right) + (1-\phi)(q_0) - F - \int_0^{\hat{\beta}} s\, ds$$

whereas social welfare under the standard treatment is simply $W_0 = q_0$.

For a given compensation δ, the expected welfare under treatment H can be rewritten as:

$$E(\mathrm{W}_H) = \phi\left(\frac{\hat{\beta}}{\bar{\beta}}\big(\alpha(q_H) + (1-\alpha)(q_0)\big) + \left(1 - \frac{\hat{\beta}}{\bar{\beta}}\right)\big(\varepsilon\alpha(q_H) + (1-\varepsilon\alpha)(q_0)\big)\right) + (1-\phi)(q_0) - F - \frac{\delta^2\bar{\beta}}{2\bar{\delta}^2}.$$

Hence, given that $\phi = \hat{\beta}/\bar{\beta} = \delta/\bar{\delta}$, with simple manipulation we find that the optimal δ solves the following problem:

$$\max_{\delta} E(\mathrm{W}_H) = \frac{\delta^2}{\bar{\delta}^2}\left(A(1-\varepsilon) - \frac{\bar{\beta}}{2}\right) + \frac{\delta}{\bar{\delta}}A\varepsilon + q_0 - F, \qquad (7)$$

where compensating patients for health data is socially optimal provided that $E(\mathrm{W}_H) \geq \mathrm{W}_0$, and an interior solution exists as long as $0 < \delta^w < \bar{\delta}$.

Proposition 2 derives the socially optimal compensation for collecting health data and determines the values of the maximum total incremental health benefit of the targeted treatment for which developing project H is welfare improving.

***Proposition* 2.** Let $\underline{A^w} = \frac{-2F(1-\varepsilon)+\sqrt{4(F(1-\varepsilon))^2+2\varepsilon^2 F\bar{\beta}}}{\varepsilon^2}$ and $\overline{A^w} = \frac{\bar{\beta}}{(2-\varepsilon)}$. *We have that:*

(i) when $\underline{A^w} \leq A < \overline{A^w}$*, the government would offer* $\delta^w = \frac{\bar{\delta}\varepsilon A}{\bar{\beta}-2A(1-\varepsilon)}$ *to patients, and the amount of collected health data for project H would be* $d^w = \frac{\varepsilon A}{\bar{\beta}-2A(1-\varepsilon)} < 1$*;*

(ii) when $A \geq \overline{A^w}$*, the government would offer* $\delta^w = \bar{\delta}$ *to patients and collect all patients' health data for project H, so as* $d^w = 1$*;*

(iii) when $A < \underline{A^w}$*, project H does not improve welfare (so as* $\delta^w = d^w = 0$*).*

Intuitively, we find that developing project H is socially optimal if the total incremental health benefit from the new treatment is high enough, that is, $A \geq \underline{A^w}$. Note that the critical value $\underline{A^w}$ is increasing in $\bar{\beta}$, meaning that higher privacy costs reduce the likelihood for targeted treatments to be socially desirable. When A is high enough (that is, higher than $\overline{A^w}$), it is socially optimal to collect health data on all patients.

6. Private incentives and social goals

We now compare private incentives to collect personal health data and invest in project H with the social optimum. Proposition 3 states that the firm may underinvest in the targeted treatment. More specifically, for intermediate values of A, investment in the new treatment would be socially optimal, but the firm prefers not to invest. Then, for higher values of A, the firm decides to invest but collects a lower amount of health data than socially desirable, thereby reducing the probability that investment in the targeted treatment is successful.

***Proposition* 3**. *Private incentives to collect health data and undertake project H are too weak. Specifically, we have that:*

(i) since $\underline{A^w} < \underline{A^H}$ *then, when* $\underline{A^w} < A < \underline{A^H}$*, the firm does not invest (i.e.,* $\delta^* = d^* = 0$*), but social welfare is higher under treatment H;*

(ii) since $\delta^* < \delta^w \leq \bar{\delta}$ then, when $\underline{A^H} \leq A < \overline{A^H}$, the amount of health data collected by the firm, and thereby the probability of success of project H, are lower than socially optimal;

(iii) it never occurs that $E(\Pi_H) > \Pi_0$ when $E(W_H) < W_0$.

The intuition for this result is that, for a given compensation δ to patients (or, equivalently, for a given amount d of collected data), data collection costs are always higher than aggregate privacy costs.[11] Formally, the firm undertakes project H as long as $E(\Pi_H) \geq \Pi_0$. For a given amount of collected data d, this condition can be rewritten as:

$$dA(1-\lambda)(d(1-\varepsilon)+\varepsilon) - \bar{\delta}d^2 - F \geq 0, \tag{8}$$

where $\widehat{\Pi}_H \equiv dA(1-\lambda)(d(1-\varepsilon)+\varepsilon)$ are expected incremental revenues from treatment H and $c(\delta) \equiv \bar{\delta}d^2$ are data collection costs. Instead, project H is socially optimal if $E(W_H) \geq W_0$, that is:

$$dA(d(1-\varepsilon)+\varepsilon) - \frac{d^2\bar{\beta}}{2} - F \geq 0, \tag{9}$$

where $\widehat{W}_H \equiv dA(d(1-\varepsilon)+\varepsilon)$ are expected incremental health benefits and $PC \equiv \frac{d^2\bar{\beta}}{2}$ are aggregate privacy costs. From Nash bargaining, it easily follows that $\widehat{W}_H > \widehat{\Pi}_H$. Hence, by comparing (8) and (9), we find that private incentives to achieve treatment H are weaker than socially desirable as long as $c(\delta) > PC$ holds. This condition can be written as $\bar{\delta}d^2 > d^2\bar{\beta}/2$, that is, $\bar{\delta} > \bar{\beta}/2$. In turn, this requires that $A < \frac{\bar{\beta}}{2(1-\varepsilon)}$, which always holds under Assumption 1.[12]

It follows that the cost for collecting patients' health data to undertake project H may be too high for the firm, and this reduces private incentives to invest in the targeted treatment relative to the social optimum. In other words, high data collection costs induce the firm to invest in treatment H for a higher expected benefit of the new treatment than socially optimal. Moreover, the compensation for patients is lower, implying that less health data are collected. We also find that $(c(\delta) - PC)$ is increasing in the degree of free riding ε.

[11] Conversely, we find that private incentives to invest in project H can be stronger than in the social optimum when the firm offers personalized compensations to patients, which significantly reduces data collection costs (see Section 7).

[12] Condition $c(\delta) > PC$ can be rewritten as $\bar{\delta}\left(\frac{\hat{\beta}(\delta)}{\bar{\beta}}\right)\left(\frac{\delta}{\bar{\delta}}\right) > \left(\frac{\hat{\beta}(\delta)}{\bar{\beta}}\right)^2\left(\frac{\bar{\beta}}{2}\right)$, that is, $\delta > \hat{\beta}(\delta)/2$ (under fulfilled expectations, we have $d = \hat{\beta}(\delta)/\bar{\beta} = \delta/\bar{\delta}$). This means that, to collect a given amount of data $d = \hat{\beta}(\delta)/\bar{\beta}$, the firm needs to offer a (uniform) compensation higher than the average privacy cost among patients sharing data (the government only cares about aggregate privacy costs because the total compensation for patients sharing data cancels out in the welfare function).

We now study how the misalignment between private incentives to invest in the targeted treatment and social goals (highlighted in Proposition 3) depends on parameters $\bar{\beta}$ and ε. Consider first the effect of privacy costs as measured by $\bar{\beta}$. Let $\Delta = (\overline{A_H} - \underline{A_w})$. We find that Δ is increasing in $\bar{\beta}$.[13] It follows that a lower value of $\bar{\beta}$ reduces the mismatch between private and social incentives to invest in the new treatment. This means that reducing privacy costs can mitigate the underinvestment problem and promote data-driven pharmaceutical R&D as well as improve social welfare.

This objective may be achieved, for instance, by providing patients with more transparent information, and by improving privacy policies. Note that these measures do affect privacy costs, but may be implemented without limiting the ability to use health data for medical innovation and basically at no cost for the firm. Health data protection is also at the heart of recent EU directives that aim at ensuring security of information processing environments. Thus, the NIS 2 Directive lays down cybersecurity risk management measures and mandates obligations to entities, including those carrying out R&D activities of medicinal products.[14]

Then, we assess how the misalignment between private and social incentives to invest depends on ε,. We find that $\frac{\partial[\underline{A^w}]}{\partial \varepsilon} < 0$. Indeed, a higher ε reduces the lower bound of the total incremental health benefit of treatment H for it to improve welfare. This is because, in an interior solution for problem (7), a higher ε enlarges the market size for treatment H, so that more patients can benefit from the new treatment.

We also find that $\frac{\partial[\overline{A^H}]}{\partial \varepsilon} > 0$. Indeed, the higher the value of ε, the higher the upper bound of the total incremental health benefit of treatment H for an interior solution for problem (6) to occur. The rationale is that, for high values of the total incremental health benefit of treatment H (close to the critical value for a corner solution $\overline{A^H}$), the fraction of the market that is affected by ε is negligible and the free-riding effect prevails, thereby reducing the firm's incentive to collect health data. This makes it less likely for the firm to collect all patients' data to achieve treatment H.

[13] Indeed, $\frac{\partial \Delta}{\partial \bar{\beta}} = \frac{2}{\varepsilon(1-\lambda)+2(1-\varepsilon)(2-\lambda)} - \frac{F}{\sqrt{4F^2(1-\varepsilon)^2+2F\bar{\beta}\varepsilon^2}} > 0$ when $\bar{\beta} > \frac{F[(2(2-\lambda)-\varepsilon(3-\lambda))^2-16(1-\varepsilon)^2]}{8\varepsilon^2}$, which is satisfied under Assumption 2, given that $\frac{2F[2-\varepsilon]}{\varepsilon} > \frac{F\left[\left(2(2-\lambda)-\varepsilon(3-\lambda)\right)^2-16(1-\varepsilon)^2\right]}{8\varepsilon^2}$.

[14] See https://www.european-health-data-space.com/NIS_2_Training.html.

It follows from above that $\frac{\partial \Delta}{\partial \varepsilon} > 0$. Thus, we can conclude that a high value of ε exacerbates the underinvestment problem.

7. Personalized compensation

In the previous section, we have found that there may be underinvestment in the targeted treatment because of excessive costs of data collection. We now investigate how this issue is related to the information on patients' types that is available to the firm. Specifically, we show that data collection costs may be significantly reduced when the firm is fully informed on patients' types. In turn, this may even lead the firm to overinvest in the new treatment.

Thus, let us assume that the firm knows the privacy type β of each patient (e.g., because it can be inferred from individuals' observable attributes, such as general health conditions or socio-economic status). In this framework, the monopolist can offer personalized compensations $\delta(\beta)$ to patients, depending on their types. In doing so, the firm can save on data collection costs relative to the case where she offers a uniform compensation.

We hereby replace assumptions 1 and 2 in Section 4 respectively with the following assumptions.

Assumption 1. $\boldsymbol{bis}$**.** Let $\alpha(q_H - q_0) < min\left\{\frac{\bar{\beta}\left((2-\lambda)-\sqrt{(2-\lambda)^2-1}\right)}{(1-\varepsilon)}, \frac{\bar{\beta}}{2(1-\varepsilon)}\right\}$.

Assumption 2. $\boldsymbol{bis}$**.** Let $\bar{\beta} > F\left(\frac{4-\varepsilon(3-\lambda)-2\lambda+\sqrt{(2-\varepsilon)(1-\lambda)(6-5\varepsilon-\lambda(2-\varepsilon))}}{\varepsilon(1-\lambda)}\right)$.

These assumptions are symmetric to those in the baseline model. Assumption 1. bis ensures concavity of both profit and welfare functions under personalized compensations. Assumption 2. bis ensures that an interior solution may exist with a personalized compensation scheme or, equivalently, excludes that there are only corner solutions where the firm either uses all patients' health data to invest in project H or does not invest at all, and the same holds for the government.

Private incentives to invest

In what follows, we solve for a fulfilled-expectation equilibrium. For this purpose, first we find the marginal patient $\hat{\beta}_p$ who shares personal data, which determines the amount of collected health data and thereby the probability of success of project H.

Let $\hat{\delta}_p \equiv \delta(\hat{\beta}_p)$ be the compensation that leaves the marginal patient sharing health data with the same net benefit as when she does not share data. Let d_p be the amount of collected health data under personalized compensations and ϕ_p^e be the corresponding expected probability of success of project H. We thus have that $\phi_p^e = d_p^e = \hat{\beta}_p/\bar{\beta}$. Then, following the same steps as in the baseline model (see equations (4) and (5)), we obtain that, in a fulfilled-expectation equilibrium, the marginal patient is indifferent between sharing health data or not when the following condition holds:

$$\hat{\beta}_p = \frac{\hat{\beta}_p}{\bar{\beta}} A(1-\varepsilon) + \hat{\delta}_p.$$

Hence, we find that:

$$\hat{\beta}_p = (\hat{\delta}_p \bar{\beta})/\bar{\delta}. \tag{10}$$

By rearranging (10), we obtain that the compensation for the marginal patient is:

$$\hat{\delta}_p = \hat{\beta}_p(\bar{\delta}/\bar{\beta}), \tag{11}$$

with $\hat{\delta}_p \leq \bar{\delta}$. Patients with privacy types β such that $\beta < \hat{\beta}_p$ who receive personalized compensations $\delta(\beta)$ decide to share health data when the expected benefit is at least as high as if they do not share data. By using $\phi_p^e = d_p^e = \hat{\beta}_p/\bar{\beta}$, from (10) we find that this occurs when:

$$\beta \leq \frac{\hat{\delta}_p}{\bar{\delta}} A(1-\varepsilon) + \delta(\beta). \tag{12}$$

Since the firm finds it profitable to extract as much surplus as possible from each patient, then she chooses the minimum (personalized) compensation $\delta(\beta)$ necessary to induce each patient with privacy type $\beta < \hat{\beta}_p$ to share health data. From equation (12), we have:

$$\delta(\beta) = \beta - \frac{\hat{\delta}_p}{\bar{\delta}} A(1-\varepsilon). \tag{13}$$

In contrast with a uniform compensation scheme, under personalized compensations each patient who shares health data does accept a compensation lower than her privacy cost. However, as with uniform compensation, (*ceteris paribus*) the level of compensation necessary to induce the marginal patient to share health data increases with ε. In the limit case where $\varepsilon = 1$, eligible patients who do not share data are perfectly identified *ex post* (e.g., via a biomarker test). This implies that the compensation for each patient matches her privacy cost.

In what follows, we assume that $\delta(\beta) \geq 0$. In doing so, we exclude that the firm may receive compensations from patients sharing health data. From (12), we can easily obtain that the patient who is indifferent between sharing health data or not with zero compensation is $\tilde{\beta} = \frac{\hat{\delta}_p}{\bar{\delta}} A(1-\varepsilon)$. Given that $\delta(\beta) \geq 0$, patients with privacy types $\beta \in [0, \tilde{\beta}]$ are willing to share health data for free.

Instead, patients with privacy types $\beta \in [\tilde{\beta}, \hat{\beta}_p]$ decide to share personal data under strictly positive compensations and thus concur to determine data collection costs. It follows that the total cost of data collection for the firm is:[15]

$$c\big(\delta(\beta)\big) = \int_{\tilde{\beta}}^{\hat{\beta}_p} \frac{1}{\bar{\beta}} c\big(\delta(s)\big) ds = \int_{\tilde{\beta}}^{\hat{\beta}_p} \frac{1}{\bar{\beta}} \left\{ s - \frac{\hat{\delta}_p}{\bar{\delta}} A(1-\varepsilon) \right\} ds = \frac{\hat{\delta}_p{}^2}{2\bar{\beta}}. \tag{14}$$

Under personalized compensations, the firm chooses $\hat{\delta}_p$ to maximize expected profit. In turn, this choice determines the amount d_p of collected health data, and thereby the probability of success ϕ_p of project *H* (where $\phi_p = \hat{\delta}_p / \bar{\delta}$). Thus, the monopolist's problem in the first stage can be written as:

$$\max_{\hat{\delta}_p} E(\Pi_H) = \frac{\hat{\delta}_p}{\bar{\delta}} A(1-\lambda) \left(\frac{\hat{\delta}_p}{\bar{\delta}} (1-\varepsilon) + \varepsilon \right) - \frac{\hat{\delta}_p{}^2}{2\bar{\beta}} + p_0 - F, \tag{15}$$

where the monopolist collects health data if and only if $E(\Pi_H) \geq \Pi_0$, and an interior solution exists as long as $0 < \hat{\delta}_p < \bar{\delta}$.

Proposition 4 defines the values of A for which the firm finds it profitable to offer personalized compensations to patients to collect health data and develop project *H*.

Proposition 4. Let $\underline{A^{Hp}} = \frac{-2\bar{\beta}F(1-\varepsilon)(2-\lambda) + \bar{\beta}\sqrt{[2F(1-\varepsilon)(2-\lambda)]^2 + 2F[\bar{\beta}\varepsilon^2(1-\lambda)^2 - 2F(1-\varepsilon)^2]}}{[\bar{\beta}\varepsilon^2(1-\lambda)^2 - 2F(1-\varepsilon)^2]}$ and $\overline{A^{Hp}} = \frac{\bar{\beta}[4-3\varepsilon-\lambda(2-\varepsilon)] - \bar{\beta}\sqrt{(1-\lambda)(2-\varepsilon)(6-5\varepsilon-2\lambda+\lambda\varepsilon)}}{2(1-\varepsilon)^2}$. *We have that:*

(i) when $\underline{A^{Hp}} \leq A < \overline{A^{Hp}}$, *the firm offers* $\delta^*(\beta) = \left(\beta - \frac{\hat{\delta}_p{}^*}{\bar{\delta}} A(1-\varepsilon) \right)$ *to patients with privacy types* $\beta \in [\tilde{\beta}(\hat{\delta}_p{}^*), \hat{\beta}_p]$, *whereas privacy types* $\beta \in [0, \tilde{\beta}(\hat{\delta}_p{}^*))$ *do share personal data for free, where* $\hat{\delta}_p{}^* =$

[15] Detailed calculations are available from the authors upon request.

$\frac{\bar{\delta}\bar{\beta}\varepsilon(1-\lambda)A}{\bar{\beta}^2+(A)^2(1-\varepsilon)^2-2\bar{\beta}A(1-\varepsilon)(2-\lambda)}$ *is the compensation offered to the marginal type sharing data* $\hat{\beta}_p$*. Thus, the amount of health data collected for project H is* $d_p^* = \hat{\delta}_p{}^*/\bar{\delta} < 1$*;*

(ii) when $A \geq \overline{A^{Hp}}$*, the firm offers* $\delta^*(\beta) = \beta - A(1-\varepsilon)$ *to patients with privacy types* $\beta \in \left[\tilde{\beta}(\bar{\delta}), \bar{\beta}\right]$*, whereas privacy types* $\beta \in [0, \tilde{\beta}(\bar{\delta}))$ *do share personal data for free. Thus, the firm collects all patients' health data for project H, that is,* $d_p^* = 1$*;*

(iii) when $A < \underline{A^{Hp}}$*, project H is not profitable.*

Consistent with the results obtained under uniform compensation (see Section 4), we find that both the critical value $\underline{A^{Hp}}$ for investment in project H in an interior solution and the critical value $\overline{A^{Hp}}$ for investment in project H using all patients' data are increasing in $\bar{\beta}$. This confirms the result that higher privacy concerns reduce the likelihood of developing the targeted treatment.

Comparison with the social optimum

In this section, we compare private incentives to invest in the targeted treatment under personalized compensations with the social optimum. It is worth noting that the compensation scheme is irrelevant from a welfare perspective. Indeed, data collection costs for the firm are exactly equal to benefits from received compensations for patients, so that these two terms cancel out in the welfare function.

Hence, the socially optimal amount of health data to be collected and the corresponding probability of success of project H, as well as the critical values of the total incremental health benefit of the new treatment for it to be socially desirable (see Section 5) are not affected by the compensation scheme.

The striking difference under personalized compensations is that, for a given compensation that leaves the marginal patient indifferent between sharing health data or not, data collection costs for the firm are lower than aggregate privacy costs for patients. Indeed, for a given value of $\hat{\delta}_p$, $\int_0^{\hat{\beta}_p} s\, ds = \frac{\hat{\delta}_p{}^2\bar{\beta}}{2\bar{\delta}^2} > \frac{\hat{\delta}_p{}^2}{2\bar{\beta}} = c(\delta)$ holds.

As shown in Section 6, this is a necessary condition to find outcomes of the game where the firm overinvests in the targeted treatment. In what follows, we show that this condition may also be sufficient for overinvestment as long as the firm has strong bargaining power in price negotiation with the government (i.e.,

λ is low enough).[16] The rationale is that the firm can reap most of the benefits from developing the targeted treatment.

Let us first consider interior solutions for problems (15) and (7) where both the firm and the government choose to collect health data from a fraction of patients. We show that in such a case the firm may collect an excessive amount of health data relative to the social optimum (see Lemma 1).

***Lemma* 1.** *Let* $0 < d_p^* < 1$ *and* $0 < d^w < 1$. *Let also* $\underline{A} = \frac{\bar{\beta} - \bar{\beta}\sqrt{1-\lambda}}{(1-\varepsilon)}$ *and* $\overline{A} = \frac{\bar{\beta} + \bar{\beta}\sqrt{1-\lambda}}{(1-\varepsilon)}$. *Under personalized compensations, we have that* $d_p^* > d^w$ *as long as* $\underline{A} < A < \overline{A}$.

Then, we can use Lemma 1 to prove the following proposition.

***Proposition* 5**. *Let* $\bar{\lambda} = \frac{3-4\varepsilon+\varepsilon^2}{4-4\varepsilon+\varepsilon^2}$. *Under personalized compensations, private incentives to collect health data and develop project H may be too strong. Specifically, this does occur when* $0 < \lambda < \bar{\lambda} \leq 3/4$, *provided that* $A \in I \subseteq \left[\underline{A^{Hp}}, \overline{A^w}\right]$.

It is worth noting that a higher degree of free riding in sharing data (i.e., a higher ε) reduces the likelihood of overinvestment (namely, $\bar{\lambda}$ is decreasing in ε). Indeed, a higher value of ε makes it more costly to collect a given amount of health data, so that the firm needs a stronger bargaining power (and thereby a higher price for the new treatment) to find it profitable to develop project H.

8. Policy implications

The results obtained in the baseline model highlight a mismatch between private and social incentives to invest in targeted treatments (see Section 6). More specifically, under uniform compensation from Proposition 3 we find underinvestment by the pharmaceutical firm.

In what follows, we discuss policy measures that may align private and social incentives for data-driven pharmaceutical innovation. We consider different options for accessing health data where the firm does not have to collect personal data directly from patients. In each of the considered cases, we assume that a central (public) agency controls health data and that patients are not compensated for sharing personal data with the agency. Centralized governance of health data is a means to avoid externalities caused by each patient's

[16] Depending on λ and A, underinvestment may occur. Detailed results are available from the authors on request.

personal data sharing decision, which is a major reason for underinvestment in the targeted treatment under uniform compensation. More specifically, in Section 8.1, the agency can provide the firm with free access to all patients' data; in Section 8.2, the firm can access all patients' data for a lump-sum payment to the agency; and finally, in Section 8.3, the agency chooses the amount of data to be made freely available to the firm.

8.1 Access to data without restrictions

Let us first consider the case where the central (public) agency can provide the firm with free access to patients' health data. Then, in the absence of data collection costs, the expected profit from the innovative treatment is:

$$E(\Pi_H) = d^2 A(1-\lambda)(1-\varepsilon) + dA\varepsilon(1-\lambda) + p_0 - F.$$

Since $E(\Pi_H)$ is increasing in the amount d of health data, then the firm will access all patients' data whenever they are freely available. Thus, profit from project H simplifies to $\Pi_H = A(1-\lambda) + p_0 - F$.[17] As long as $\Pi_H \geq \Pi_0$, the firm finds it profitable to develop project H by using all patients' data. This occurs when $A \geq \frac{F}{(1-\lambda)}$.[18] Since we have that $\frac{F}{(1-\lambda)} < \underline{A_H}$ then, as expected, free access to health data increases the likelihood of private investment in the new treatment, by reducing the critical value of the total incremental health benefit for investment to occur relative to the baseline model.

We now compare private incentives to invest in the targeted treatment under this policy option with the social optimum. First, note that the socially optimal investment determined in Section 5 is not affected by the policy at issue, because total costs of data collection for the firm and the aggregate compensation for patients cancel out in the welfare function. Therefore, social welfare maximization leads to d^w (see Proposition 2).

We find that providing the firm with free access to health data may yield overinvestment in the targeted treatment when the firm's bargaining power is sufficiently high. More specifically, the firm may find it profitable to invest by using all patients' data when, from a welfare perspective, it would be optimal to invest by using data from a fraction of patients, or not to invest at all. Intuitively, high privacy costs (i.e., a high value of $\bar{\beta}$) are more likely to induce overinvestment. Indeed, high privacy costs do not affect private incentives to

[17] When all patients' health data are used (i.e., $d = 1$), Π_H does not depend on ε.

[18] Instead, when $A < \frac{F}{(1-\lambda)}$ investment in treatment H is totally precluded, since the firm does not invest even if she has free access to all patient's health data.

invest (the critical value for investment $\frac{F}{(1-\lambda)}$ does not depend on $\bar{\beta}$), whereas they reduce social incentives to invest. Proposition 6 summarizes the result.

***Proposition* 6.** *A policy granting free access to health data may lead to overinvestment in the new treatment. Specifically, if the firm's bargaining power is high enough, that is,* $\lambda < \frac{\bar{\beta}-F(2-\varepsilon)}{\bar{\beta}} \equiv \lambda_2$ *then the firm overinvests when* $\frac{F}{(1-\lambda)} < A < \overline{A^W}$ *(otherwise, the firm may underinvest). Overinvestment is more likely when privacy costs are high (that is,* λ_2 *is increasing in* $\bar{\beta}$*).*

We now study whether and when this policy improves social welfare relative to the baseline model. We point out that there are two intervals of values of the total incremental health benefit of treatment H where investment decisions are misaligned. When $\frac{F}{(1-\lambda)} < A < \underline{A^H}$, the firm invests only when she has free access to all patients' data. Instead, when $\underline{A^H} < A < \overline{A^H}$, in the baseline model the firm invests by collecting a fraction of patients' health data.

Proposition 7 states that social welfare is higher (respectively, lower) when the firm has free access to all patients' data than in the baseline model as long as the total incremental health benefit of the targeted treatment is sufficiently large (respectively, small). Moreover, free access to health data is less likely to improve welfare when patients' privacy costs are high.

***Proposition* 7.** *Let* $\beta_l = \frac{2F\lambda}{1-\lambda}$, $\beta_h = \frac{2F\left(\varepsilon\left(2(2-\lambda)-\varepsilon(1-\lambda)^2\right)+2\lambda\right)+4F\sqrt{2\varepsilon^2+4\varepsilon\lambda+\lambda^2-\varepsilon(2+\varepsilon)\lambda^2}}{\varepsilon^2(1-\lambda)^2}$, *and* $\tilde{A} = \frac{\bar{\beta}(8-5\varepsilon-2\lambda+3\varepsilon\lambda)-\bar{\beta}\sqrt{9\varepsilon^2+2(10-7\varepsilon)\varepsilon\lambda+(2-3\varepsilon)^2\lambda^2}}{4(1-\varepsilon)(4-\varepsilon(1-\lambda)-2\lambda)}$. *A policy granting the firm free access to health data can improve social welfare. Specifically, we have that:*

(i) when $max\left\{\frac{F}{(1-\lambda)}, \frac{\bar{\beta}}{2}+F\right\} < A < \underline{A^H}$*, developing project H with free access to all patients' data improves welfare relative to the baseline model where there is no investment, given that* $\bar{\beta} < \beta_h$*;*

(ii) when $max\{\underline{A^H}, \tilde{A}\} < A < \overline{A^H}$*, developing project H with free access to all data improves welfare relative to the baseline model where the firm collects only a fraction of patients' data. For sufficiently low values of* $\bar{\beta}$*, namely* $\bar{\beta} < \beta_l$ *(with* $\beta_l < \beta_h$*), the policy never reduces welfare.*

We conclude that, compared to the baseline model, social welfare can either increase or decrease when the firm can freely access all patients' health data, depending on the total incremental health benefit of the targeted

treatment and on patients' privacy costs. More specifically, a sufficiently low value of $\bar{\beta}$ ensures that the policy at issue improves social welfare whenever it induces a different investment choice than in the baseline model.

8.2 Lump-sum payment for access to health data

In this section, we assume that the central (public) agency that manages patients' health data provides the firm with access to such data at a price $P \geq 0$. Thus, the firm makes a lump-sum payment to the agency for a centralized access to health data, and then chooses the amount of data to be used for investing in project H. The timing of the game is as follows.

1. The agency sets the price P for access to patients' health data.
2. The monopolist observes the price P and decides whether to pay to access health data and undertake project H. Under project H, the firm chooses the amount of data to use for research.
3. If the targeted treatment is achieved, then the firm and the government negotiate the price of the treatment through a (generalized) Nash bargaining process.

Stage 3 proceeds as usual. Then, for a given anticipated price p_H for the targeted treatment, in stage 2 the expected profit under project H can be written as:

$$E(\Pi_H) = d^2 A(1-\lambda)(1-\varepsilon) + dA\varepsilon(1-\lambda) + p_0 - P - F,$$

with $0 \leq d \leq 1$. Because the expected profit is increasing in d, the firm prefers to access all data for a given payment P, so that $d = 1$. Therefore, when deciding whether to pay for access to health data, the firm compares the expected profit from project H by investing with all patients' data, that is, $E(\Pi_H) = A(1-\lambda) + p_0 - P - F$, with profit in the *status quo*, that is, $\Pi_0 = p_0$. Hence, the firm decides to pay P and undertake project H with all patients' data if:

$$P \leq A(1-\lambda) - F = \bar{P}.$$

This implies that, for any $P \geq 0$, the firm does not invest when $A < \frac{F}{(1-\lambda)}$. In stage 1, the agency anticipates that the firm uses all patients' data under project H, i.e., $d = 1$. Thus, the expected welfare with investment in treatment H is $E(W_H) = A + q_0 - \frac{\bar{\beta}}{2} - F$.[19] The agency chooses P to maximize welfare, so that it induces the

[19] Indeed, the lump-sum payment P cancels out in the welfare function.

firm to pay for access to data and invest in project H (because $E(\Pi_H) \geq \Pi_0$ holds with $d = 1$) when $E(W_H) \geq W_0$ also holds with $d = 1$. This occurs when $A \geq \frac{\bar{\beta}}{2} + F$.

Inserting for $A = \frac{\bar{\beta}}{2} + F$ in $\bar{P}$, we find $\bar{P} = \frac{\bar{\beta}}{2}(1 - \lambda) - \lambda F$, that is non-negative when $\frac{F}{(1-\lambda)} \leq \frac{\bar{\beta}}{2} + F$. Given that $d = 1$, this price for access to data is welfare maximizing because it implies that the firm finds it profitable to incur the lump-sum payment and invest in the targeted treatment exactly when $A \geq \frac{\bar{\beta}}{2} + F \geq \frac{F}{(1-\lambda)}$ holds. Thus, when all patients' health data are used for project H, this policy can align private incentives to invest in treatment H with social goals.[20] Nonetheless, the first-best outcome is such that, when $\underline{A^w} < A < \overline{A^w}$, investment in treatment H should occur with $d^w < 1$. By comparing private and social incentives to invest, we find that the proposed policy leads to underinvestment when $\underline{A^w} < A < \frac{\bar{\beta}}{2} + F$, and to overinvestment when $\frac{\bar{\beta}}{2} + F < A < \overline{A^w}$.

Compared to the baseline model, this policy can be welfare improving as long as the total incremental health benefit of the targeted treatment is sufficiently high. Note that the conditions for a welfare improvement are the same as in Proposition 7. However, the lump-sum payment for health data avoids the excessive data use registered under free access to data when the firm's bargaining power or privacy costs are high, which reduces welfare. Specifically, the lump-sum payment for patients' data prevents data use as long as $\frac{F}{(1-\lambda)} < A < \frac{\bar{\beta}}{2} + F$, where social welfare is higher without investment than with investment in H using all patients' data (that occurs under free access to data).

8.3 Restricted access to health data

In this section, we consider the case where the central (public) agency grants the firm free access to the amount of data that maximizes social welfare under project H, provided that this improves welfare relative to the *status quo* (that is, $E(\mathrm{W}_H) \geq \mathrm{W}_0$) and that such amount of data is large enough to induce the firm to invest in the targeted treatment (that is, $E(\Pi_H) \geq \Pi_0$). Given these constraints, the public agency's problem is:

[20] If the firm has low bargaining power (namely, $\lambda > \frac{\bar{\beta}}{2F+\bar{\beta}}$), then $\frac{\bar{\beta}}{2} + F < \frac{F}{(1-\lambda)}$ and the policy cannot induce investment when $\frac{\bar{\beta}}{2} + F < A < \frac{F}{(1-\lambda)}$.

$$\max_d E(\mathrm{W}_H) = d\alpha(q_H - q_0)(d(1-\varepsilon)+\varepsilon) + q_0 - d^2\frac{\bar{\beta}}{2} - F,$$

with $0 \leq d \leq 1$. From $E(\Pi_H) \geq \Pi_0$, namely $d^2A(1-\lambda)(1-\varepsilon) + dA\varepsilon(1-\lambda) - F \geq 0$, we derive a lower bound d^{min} for the amount of health data that is necessary for the firm to invest in project H. Specifically, we obtain that the firm finds it profitable to invest if and only if $d \geq \frac{-A\varepsilon(1-\lambda)+\sqrt{(A)^2\varepsilon^2(1-\lambda)^2+4FA(1-\lambda)(1-\varepsilon)}}{2A(1-\lambda)(1-\varepsilon)} \equiv d^{min}$. Thus, when $d^w \geq d^{min}$, solving the public agency's problem leads to achieve the first-best outcome (see Proposition 2).[21]

We show that this occurs when the firm has enough bargaining power (i.e., λ is sufficiently low) or when the firm has less bargaining power (i.e., λ is higher) and the total incremental health benefit of the new treatment is sufficiently high. If this benefit takes intermediate values, then the profit constraint is binding in equilibrium, and the first best cannot be achieved. Thus, the agency decides whether to grant free access to a larger amount of patients' data than the socially optimal amount thereby inducing the firm to invest in the targeted treatment. In such a case, social welfare increases relative to the baseline model. Finally, when the firm has weak bargaining power (i.e., λ is high enough), the first best is achieved by granting the firm access to all patients' data only if the total incremental health benefit of the new treatment is very high (otherwise, the policy does not improve welfare). Proposition 8 summarizes the results.

***Proposition* 8.** *Let $\hat{A}$ be such that $(d^{min}|A=\hat{A}) = (d^w|A=\hat{A})$ and $\bar{\bar{A}}$ be such that $E\left(W_H|d=d^{min}(\bar{\bar{A}})\right) = W_0$. Then, a policy granting free access to an amount of health data selected by a public agency mitigates the underinvestment problem and may improve social welfare, depending on the firm's bargaining power. More specifically:*

i) if $\lambda \leq \frac{1}{2} + \sqrt{\frac{F(1-\varepsilon)^2}{4F(1-\varepsilon)^2+2\bar{\beta}\varepsilon^2}} \equiv \lambda_1$ then the policy achieves first best;

ii) if $\lambda_1 < \lambda < \frac{\bar{\beta}-F(2-\varepsilon)}{\bar{\beta}} \equiv \lambda_2$ then the policy achieves first best when $A \geq \hat{A}$, with $\hat{A} \in (\underline{A^w}, \overline{A^w})$. When $\underline{A^w} < \bar{\bar{A}} < A < \hat{A}$, the policy induces a welfare-improving investment in the new treatment by providing the firm with an amount of data $d^{min} > d^w$;

iii) if $\lambda \geq \lambda_2$ then the policy achieves first best when $A \geq \frac{F}{(1-\lambda)}$, otherwise it is not applied.

[21] Recall that, since $d^w = \delta^w/\bar{\delta}$, then maximizing social welfare with respect to d or δ leads to an equivalent outcome.

To summarize, Proposition 8 finds that the proposed policy mitigates the underinvestment problem because it expands both the range of values of the total incremental health benefit of the targeted treatment for which investment occurs and the amount of health data used to develop the new treatment. Given that there is investment in the new treatment, social welfare is higher under the policy at issue than in the baseline model.

9 Conclusions

We have studied private incentives to invest in a targeted treatment for an eligible patient group based on collected personal health data, and compared them with the social optimum. In our paper, patients must be compensated for sharing health data, because they incur privacy costs and can partially free ride on the public good nature of data-driven drug innovation. In the baseline model, we have found that the pharmaceutical firm underinvests in the targeted treatment. We have shown that this occurs because (for a given compensation to patients) total costs of health data collection for the firm are higher than aggregate privacy costs. Therefore, privacy protection measures reducing patients' costs of sharing health data can contribute to align private investment incentives with social goals.

On the other hand, we have shown that the firm may overinvest when she is fully informed about each patient's privacy cost, so that she can offer a personalized (instead of a uniform) compensation to patients for sharing health data. In this case, for a given compensation to the marginal patient sharing data, data collection costs are lower than aggregate privacy costs.

Overinvestment in the new treatment may also take place in an alternative scenario where health data are governed by a central (public) agency and patients are not compensated for sharing data. This particularly occurs if the firm is provided with free access to health data for medical research, when the firm's bargaining power in price negotiation is strong enough and with high privacy costs. If the firm has to make a lump-sum payment to the agency to access health data, then overinvestment can be mitigated. Compared to the baseline model, both policies improve social welfare when the total incremental health benefit of the new treatment is high enough (especially if privacy costs are low).

Finally, a policy granting free access to an amount of health data selected by the agency achieves the first-best outcome when the firm has enough bargaining power, or when the total incremental health benefit of the new treatment is sufficiently high. If such benefit takes intermediate values, then the agency grants free access

to a larger amount of health data than socially optimal to foster investment in the new treatment. In turn, this improves social welfare relative to the baseline model.

Although it seems very promising, the process of integration of AI tools in health care is far from being complete, and the ultimate success of this approach to disease treatment and prevention depends on many factors (*in primis*, easy access to patients' health data) as well as on the interplay among private organizations (pharmaceutical firms and R&D laboratories) and public bodies (governments, health authorities, and regulatory agencies), whose incentives and goals are not necessarily aligned.

In future work, we may depart from the assumption of full reimbursement of drugs and/or drug price negotiation to study whether data-driven pharmaceutical R&D widens social disparities when it implies high treatment prices for patients. We may also examine an alternative framework where two independent R&D laboratories are engaged to develop the targeted treatment, so as to investigate whether and when (potential) competition may improve data-driven pharmaceutical innovation.

Acknowledgements

The authors thank Philippe De Donder, Yassine Lefouili, Izabela Jelovac, and the TEPREME project research team for helpful comments and discussions, as well as seminar audiences at Toulouse (DHW and TBS 2024), Sevilla (JEI 2024), Bergamo (SIEPI 2024), and Roma (EARIE 2023). The authors gratefully acknowledge funding from Agence Nationale de la Recherche (ANR) under the TEPREME project (ANR–21–CE26–0014–01). Anna D'Annunzio also acknowledges funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program). Chiara Conti and Pierfrancesco Reverberi also acknowledge funding from Sapienza University of Rome under grant RP122181681C581A.

References

Antoñanzas, F., Juárez-Castelló, C.A., Rodríguez-Ibeas, R., 2015. Some economics on personalized and predictive medicine. *European Journal of Health Economics*, 16, 985–994.

Bardey, D., Bommier, A., Jullien, B., 2010. Retail price regulation and innovation: Reference pricing in the pharmaceuticals industry. *Journal of Health Economics*, 29, 303–316.

Barigozzi, F., Jelovac, I., 2020. Research funding and price negotiation for new drugs, *Health Economics*, 29, 83–96.

Brekke, K.R., Dalen, D.M., Straume, O.R., 2024. Competing with precision: Incentives for developing predictive biomarker tests. *The Scandinavian Journal of Economics*, 126, 60-97.

Brekke, K.R., Dalen, D.M., Straume, O.R., 2023. The price of cost-effectiveness thresholds under therapeutic competition in pharmaceutical markets. *Journal of Health Economics*, 90, 102778.

Canta, C., D'Annunzio, A., Lefouili, Y., 2025. Data sharing, privacy, and innovation in healthcare. Mimeo.

Conti, C., D'Annunzio, A., Reverberi, P., 2024. Data-driven health innovation and privacy regulation. *The BE Journal of Economic Analysis & Policy* 24.1, 329–338.

Conti, C., Reverberi, P., 2021. Price discrimination and product quality under opt-in privacy regulation. *Information Economics and Policy*, 55, 100912.

EC–European Commission, 2022. A European health data space: Harnessing the power of health data for people, patients and innovation. COM(2022), 196 final.

Gavan, F.P., Thompson, A.J., Payne, K., 2018. The economic case for precision medicine. *Expert Review of Precision Medicine and Drug Development*, 3, 1-9.

Gonzalez, P., Macho-Stadler, I., Perez-Castrillo, D., 2016. Private versus social incentives for pharmaceutical innovation. *Journal of Health Economics*, 50, 286–297.

Lefouili, Y., Madio, L., Toh, Y.L., 2024. Privacy regulation and quality-enhancing innovation, *The Journal of Industrial Economics*, 0022-1821.

Matteucci, G., Reverberi, P. 2017. Drug innovation, price controls, and parallel trade, *International Journal of Health Economics and Management*, 17, 159–179.

Miller, A.R., Tucker, C., 2018. Privacy protection, personalized medicine and genetic testing. *Management Science*, 64, 4648–4668.

Mougeot, M., Naegelen, F., 2022. Incentives to implement personalized medicine under second-best pricing. *Health Economics*, DOI: 10.1002/hec.4588.

Rhodes, A., Zhou, J., 2024. Personalized pricing and competition. *American Economic Review*, 114, 2141–2170.

Stern, A.D., Alexander, B.M., Chandra, A., 2017. How economics can shape precision medicines. *Science*, 355, 1131–1133.

World Health Organization (WHO), 2021. Precision Medicine: Opportunities and Challenges.

Appendix

Proof of Proposition 1.

From eq. (6), the expected profit under investment in project H is:

$$E(\Pi_H) = \frac{\delta}{\bar{\delta}}(1-\lambda)A\left(\frac{\delta}{\bar{\delta}}(1-\varepsilon)+\varepsilon\right) - \frac{\delta^2}{\bar{\delta}} + p_0 - F$$

with $\bar{\delta} = \bar{\beta} - A(1-\varepsilon)$. Rearranging the above formula, we obtain:

$$E(\Pi_H) = \frac{\delta^2}{\bar{\delta}^2}\left(A(2-\lambda)(1-\varepsilon) - \bar{\beta}\right) + \frac{\delta}{\bar{\delta}}A(1-\lambda)\varepsilon + p_0 - F.$$

Then, the first order condition (FOC) on $E(\Pi_H)$ yields:[22]

$$\delta^* = \frac{\bar{\delta}\varepsilon(1-\lambda)A}{2\left(\bar{\beta} - A(2-\lambda)(1-\varepsilon)\right)}.$$

The firm undertakes project H as long as $E(\Pi_H) \geq \Pi_0$. Solving for $[E(\Pi_H) - \Pi_0] = \frac{\delta^2}{\bar{\delta}^2}\left(A(2-\lambda)(1-\varepsilon) - \bar{\beta}\right) + \frac{\delta}{\bar{\delta}}A(1-\lambda)\varepsilon - F = 0$ with respect to δ, we find the following roots:

$$\delta_1, \delta_2 = \frac{\bar{\delta}A\varepsilon(1-\lambda) \mp \bar{\delta}\sqrt{(A)^2\varepsilon^2(1-\lambda)^2 - 4\left(\bar{\beta} - A(2-\lambda)(1-\varepsilon)\right)}}{2\left(\bar{\beta} - A(2-\lambda)(1-\varepsilon)\right)}.$$

If the argument under square root is non-negative, then it does exist a compensation δ such that $\delta_1 < \delta < \delta_2$ for which the firm finds it profitable to invest in project H. Since $A > 0$, then the square root in the equation above is non-negative for $A \geq \frac{-2F(1-\varepsilon)(2-\lambda)+2\sqrt{[F(1-\varepsilon)(2-\lambda)]^2+F\bar{\beta}\varepsilon^2(1-\lambda)^2}}{\varepsilon^2(1-\lambda)^2} \equiv \underline{A^H}$. If $A < \underline{A^H}$, then the firm does not invest in project H. If $A \geq \underline{A^H}$, then $\delta^* = \frac{\bar{\delta}\varepsilon(1-\lambda)A}{2\left(\bar{\beta}-A(2-\lambda)(1-\varepsilon)\right)}$ maximizes the expected profit from project H as long as $0 < \frac{\bar{\delta}\varepsilon(1-\lambda)A}{2\left(\bar{\beta}-A(2-\lambda)(1-\varepsilon)\right)} < \bar{\beta} - (1-\varepsilon)A = \bar{\delta}$ holds. This occurs when $A < \frac{2\bar{\beta}}{[\varepsilon(1-\lambda)+2(1-\varepsilon)(2-\lambda)]} \equiv$

[22] The second order condition (SOC), namely, $A(2-\lambda)(1-\varepsilon) - \bar{\beta} < 0$, holds under Assumption 1.

$\overline{A^H}$. Then, there is an interior solution for δ (i.e. $\delta^* < 1$) if and only if $\underline{A^H} \leq A < \overline{A^H}$, given that $\overline{A^H} > \underline{A^H}$. We find that $\overline{A^H} > \underline{A^H}$ holds under Assumption 2.

If $A \geq \overline{A^H}$, then the compensation that maximizes the expected profit from project H is $\delta^* = \bar{\delta}$, and the firm collects all patients' health data. Let us check that, for these values of A, investment in project H is profitable. Condition $E\left(\Pi_H \middle| \delta = \bar{\delta}\right) \geq \Pi_0$ can be written as $A(2-\lambda)(1-\varepsilon) - \bar{\beta} + A(1-\lambda)\varepsilon \geq F$, or:

$$A \geq \frac{F + \bar{\beta}}{[(2-\lambda)(1-\varepsilon) + (1-\lambda)\varepsilon]}.$$

Since, under Assumption 2, $\frac{F+\bar{\beta}}{[(2-\lambda)(1-\varepsilon)+(1-\lambda)\varepsilon]} < \overline{A^H}$ holds, then when $A \geq \overline{A^H}$ the monopolist undertakes project H by collecting all patients' data. ■

***Proof of Proposition* 2.**

The FOC on $E(\mathrm{W}_H)$ yields $\delta^w = \frac{\bar{\delta}\varepsilon A}{\bar{\beta} - 2A(1-\varepsilon)}$.[23] The government would undertake project H as long as $E(\mathrm{W}_H) \geq \mathrm{W}_0$. Solving for $[E(\mathrm{W}_H) - \mathrm{W}_0] = 0$ with respect to δ, we find the following roots:

$$\delta_3, \delta_4 = \frac{\bar{\delta}A\varepsilon \mp \bar{\delta}\sqrt{(A)^2\varepsilon^2 - 2F(\bar{\beta} - 2A(1-\varepsilon))}}{(\bar{\beta} - 2A(1-\varepsilon))}.$$

If the argument under square root is non-negative, then it does exist a compensation δ such that $\delta_3 < \delta < \delta_4$ for which the government would invest in project H. Since $A > 0$, then this condition holds for $A \geq \frac{-2F(1-\varepsilon)+\sqrt{4[F(1-\varepsilon)]^2+2\varepsilon^2 F\bar{\beta}}}{\varepsilon^2} \equiv \underline{A^w}$. If $A < \underline{A^w}$, then the government would not invest in project H. If $A \geq \underline{A^w}$, then $\delta^w = \frac{\bar{\delta}\varepsilon A}{\bar{\beta} - 2A(1-\varepsilon)}$ maximizes the expected welfare under project H as long as $0 \leq \delta^w \leq \bar{\delta}$ holds. This occurs when $A \leq \frac{\bar{\beta}}{(2-\varepsilon)} \equiv \overline{A^w}$. Then, there is an interior solution for δ if and only if $\underline{A^w} < A < \overline{A^w}$, given that $\overline{A^w} > \underline{A^w}$. We find that $\overline{A^w} > \underline{A^w}$ holds under Assumption 2.

If $A \geq \overline{A^w}$, then the compensation that maximizes the expected welfare under project H is $\delta^w = \bar{\delta}$, and the government would collect all patients' health data. Let us check that, for these values of A, investment in project H does improve welfare. Condition $E(\mathrm{W}_H | \delta = \bar{\delta}) \geq \mathrm{W}_0$ implies that $A > F + \frac{\bar{\beta}}{2}$ must hold. Since $F +$

[23] The SOC, namely, $A < \bar{\beta}/(2(1-\varepsilon))$, holds under Assumption 1.

$\frac{\bar{\beta}}{2} < \overline{A^w}$ holds under Assumption 2, then collecting all patients' health data to develop project H is socially optimal when $A \geq \overline{A^w}$. ■

***Proof of Proposition* 3.**

First, we show that $\underline{A^w} < \underline{A^H}$. This inequality can be written as:

$$2F(1-\varepsilon)[(2-\lambda)-(1-\lambda)^2] + (1-\lambda)^2\sqrt{4\big(F(1-\varepsilon)\big)^2 + 2(\varepsilon)^2F\bar{\beta}}$$

$$< 2\sqrt{F^2(2-\lambda)^2(1-\varepsilon)^2 + \varepsilon^2(1-\lambda)^2F\bar{\beta}}$$

Then, squaring both sides, we obtain:

$$4F^2(1-\varepsilon)^2[(2-\lambda)^2 + (1-\lambda)^4 - 2(2-\lambda)(1-\lambda)^2] + (1-\lambda)^4\left[4\big(F(1-\varepsilon)\big)^2 + 2(\varepsilon)^2F\bar{\beta}\right]$$

$$+ 4F(1-\varepsilon)[(2-\lambda)-(1-\lambda)^2](1-\lambda)^2\sqrt{4\big(F(1-\varepsilon)\big)^2 + 2(\varepsilon)^2F\bar{\beta}}$$

$$< 4[F^2(2-\lambda)^2(1-\varepsilon)^2 + \varepsilon^2(1-\lambda)^2F\bar{\beta}]$$

or, equivalently:

$$2(1-\varepsilon)[(2-\lambda)-(1-\lambda)^2]\sqrt{4\big(F(1-\varepsilon)\big)^2 + 2(\varepsilon)^2F\bar{\beta}}$$

$$< \varepsilon^2\bar{\beta}[2-(1-\lambda)^2] + 4F(1-\varepsilon)^2[(2-\lambda)-(1-\lambda)^2]$$

Squaring again both sides, the inequality becomes:

$$4(1-\varepsilon)^2[(2-\lambda)^2 + (1-\lambda)^4 - 2(2-\lambda)(1-\lambda)^2]\left[4\big(F(1-\varepsilon)\big)^2 + 2(\varepsilon)^2F\bar{\beta}\right]$$

$$< \varepsilon^4\bar{\beta}^2[4 + (1-\lambda)^4 - 4(1-\lambda)^2]$$

$$+ 16F^2(1-\varepsilon)^4[(2-\lambda)^2 + (1-\lambda)^4 - 2(2-\lambda)(1-\lambda)^2]$$

$$+ 8F(1-\varepsilon)^2[(2-\lambda)-(1-\lambda)^2]\varepsilon^2\bar{\beta}[2-(1-\lambda)^2]$$

Then, rearranging the terms and further simplifying, we obtain:

$$\lambda[8F(1-\varepsilon)^2][(1-\lambda)^2 - (2-\lambda)] < \varepsilon^2\bar{\beta}[2-(1-\lambda)^2]^2$$

which can be expressed as a quadratic function in ε, that is:

$$\varepsilon^2\{\bar{\beta}[2-(1-\lambda)^2]^2 + \lambda 8F[(2-\lambda)-(1-\lambda)^2]\} - 16\varepsilon\lambda F[(2-\lambda)-(1-\lambda)^2]$$

$$+ \lambda 8F[(2-\lambda)-(1-\lambda)^2] > 0$$

This condition holds because the quadratic function in ε always takes positive values. This is because the coefficient of ε^2 is positive and the function has no real roots, given that:

$$16^2\lambda^2F^2[(2-\lambda)-(1-\lambda)^2]^2$$
$$-32\{\bar{\beta}[2-(1-\lambda)^2]^2+\lambda 8F[(2-\lambda)-(1-\lambda)^2]\}\lambda F[(2-\lambda)-(1-\lambda)^2]<0$$

or equivalently:

$$-\bar{\beta}[2-(1-\lambda)^2]^2\lambda F[(2-\lambda)-(1-\lambda)^2]<0.$$

Having proved that $\underline{A^w}<\underline{A^H}$, when $\underline{A^w}<A<\underline{A^H}$ we have underinvestment in the new treatment. Indeed, the government would develop project H (since $E(W_H)>W_0$), whereas the firm would only invest in project H if $A\geq\underline{A^H}$ (when $E(\Pi_H)\geq\Pi_0$).

Let us now consider the ordering of private and social critical values for investment in project H. From Assumption 2, we have that $\underline{A^w}<\overline{A^w}$ and $\underline{A^H}<\overline{A^H}$. It can also be easily checked that $\overline{A^w}<\overline{A^H}$. It follows that there are two alternative orderings of such critical values:

a) $\underline{A^w}<\underline{A^H}<\overline{A^w}<\overline{A^H}$

b) $\underline{A^w}<\overline{A^w}<\underline{A^H}<\overline{A^H}$

Consider first case a). In this case, when $\underline{A^H}<A<\overline{A^w}$ there is an interior solution where both the firm and the government would invest in the new treatment by collecting health data from a fraction of patients (with lowest privacy costs). However, the firm offers a lower compensation to such patients, and thereby collects less health data than socially optimal. Indeed, we have that $\delta^*<\delta^w<1$ when $\frac{\varepsilon(1-\lambda)A}{2(\bar{\beta}-A(2-\lambda)(1-\varepsilon))}<\frac{\varepsilon A}{\bar{\beta}-2\alpha A(1-\varepsilon)}$, or equivalently $A<\frac{\bar{\beta}(1+\lambda)}{2(1-\varepsilon)}$. Since $\frac{\bar{\beta}(1+\lambda)}{2(1-\varepsilon)}>\frac{\bar{\beta}}{(2-\varepsilon)}=\overline{A^w}$, then $\delta^*<\delta^W<1$ always holds in the relevant range.

Instead, when $\overline{A^w}<A<\overline{A^H}$ the firm invests in treatment H by collecting health data from patients with lowest privacy costs (interior solution), whereas the government would collect all patients' data by offering $\delta=\bar{\delta}>\delta^*$ (corner solution). The same reasoning applies under case b) when $\underline{A^H}<A<\overline{A^H}$. ■

***Proof of Proposition* 4.**

From (15), the FOC on $E(\Pi_H)$ yields:

$$\hat{\delta}_p^{\,*} = \frac{\bar{\delta}\bar{\beta}\varepsilon(1-\lambda)A}{\bar{\beta}^2 + (A)^2(1-\varepsilon)^2 - 2\bar{\beta}A(1-\varepsilon)(2-\lambda)}$$

and the SOC (i.e., $A < \frac{\bar{\beta}[(2-\lambda)-\sqrt{(2-\lambda)^2-1}]}{(1-\varepsilon)}$) holds under Assumption $1.bis$. The firm carries out project H as long as $E(\Pi_H) \geq \Pi_0$. We find that $[E(\Pi_H) - \Pi_0] = \hat{\delta}_p^{\,2}\left(\frac{A(1-\lambda)(1-\varepsilon)}{\bar{\delta}^2} - \frac{1}{2\bar{\beta}}\right) + \frac{\hat{\delta}_p}{\bar{\delta}}A\varepsilon(1-\lambda) - F$, where $\left(\frac{A(1-\lambda)(1-\varepsilon)}{\bar{\delta}^2} - \frac{1}{2\bar{\beta}}\right) < 0$ holds under Assumption $1.bis$.

Solving for $[E(\Pi_H) - \Pi_0] = 0$ with respect to δ, we find the following roots:

$$\delta_{1p}, \delta_{2p} = \frac{A\bar{\beta}\bar{\delta}\varepsilon(1-\lambda) \mp \bar{\delta}\sqrt{(A)^2\bar{\beta}^2\varepsilon^2(1-\lambda)^2 - 2F\bar{\beta}[\bar{\delta}^2 - 2\bar{\beta}A(1-\lambda)(1-\varepsilon)]}}{(\bar{\delta}^2 - 2\bar{\beta}A(1-\lambda)(1-\varepsilon))}$$

If the argument under square root is non-negative, then it does exist a compensation $\hat{\delta}_p$ such that $\delta_{1p} \leq \hat{\delta}_p \leq \delta_{2p}$ for which the firm finds it profitable to develop project H. This condition is satisfied for $A \geq \frac{-2\bar{\beta}F(1-\varepsilon)(2-\lambda)+\bar{\beta}\sqrt{[2F(1-\varepsilon)(2-\lambda)]^2+2F[\bar{\beta}\varepsilon^2(1-\lambda)^2-2F(1-\varepsilon)^2]}}{[\bar{\beta}\varepsilon^2(1-\lambda)^2-2F(1-\varepsilon)^2]} \equiv \underline{A^{Hp}}$. If $A < \underline{A^{Hp}}$ then the firm does not develop project H. If $A \geq \underline{A^{Hp}}$ then $\hat{\delta}_p^{\,*}$ maximizes expected profit from project H as long as $0 \leq \hat{\delta}_p^{\,*} \leq \bar{\delta}$ holds. This occurs when $A \leq \frac{\bar{\beta}[4-3\varepsilon-\lambda(2-\varepsilon)]-\bar{\beta}\sqrt{(1-\lambda)(2-\varepsilon)(6-5\varepsilon-2\lambda+\lambda\varepsilon)}}{2(1-\varepsilon)^2} \equiv \overline{A^{Hp}}$. Then, there is an interior solution for $\hat{\delta}_p$ (i.e., $\hat{\delta}_p^{\,*} < \bar{\delta}$) if and only if $\underline{A^{Hp}} \leq A < \overline{A^{Hp}}$ (Assumption $2.bis$ ensures that $\overline{A^{Hp}} > \underline{A^{Hp}}$). From (10) and (13) we easily find that, in an interior solution, the firm offers $\delta_\beta = \beta - \frac{\hat{\delta}_p^{\,*}}{\bar{\delta}}A(1-\varepsilon) = \beta - \frac{\bar{\beta}\varepsilon(1-\varepsilon)(1-\lambda)(A)^2}{\bar{\beta}^2+(A)^2(1-\varepsilon)^2-2\bar{\beta}A(1-\varepsilon)(2-\lambda)}$ to patients with privacy types $\beta \in [\tilde{\beta}, \hat{\beta}_p]$, where $\tilde{\beta}(\hat{\delta}_p^{\,*}) = \frac{\hat{\delta}_p^{\,*}}{\bar{\delta}}A(1-\varepsilon) = \frac{\bar{\beta}\varepsilon(1-\varepsilon)(1-\lambda)(A)^2}{\bar{\beta}^2+(A)^2(1-\varepsilon)^2-2\bar{\beta}A)(1-\varepsilon)(2-\lambda)}$ and $\hat{\beta}_p(\hat{\delta}_p^{\,*}) = \frac{\bar{\beta}\hat{\delta}_p^{\,**}}{\bar{\delta}} = \frac{\bar{\beta}^2\varepsilon(1-\lambda)A}{\bar{\beta}^2+(A)^2(1-\varepsilon)^2-2\bar{\beta}A(1-\varepsilon)(2-\lambda)} < \bar{\beta}$. The amount of collected data is $d_p^* = \frac{\hat{\delta}_p^{\,*}}{\bar{\delta}} < 1$.

If $A \geq \overline{A^{Hp}}$, then the compensation to the marginal patient that maximizes expected profit from project H is $\hat{\delta}_p^{\,*} = \bar{\delta}$ and the firm collects all patients' health data. Condition $E(\Pi_H|\delta = \bar{\delta}) \geq \Pi_0$ can be written as $A(1-\lambda) - \frac{[\bar{\beta}^2+(A)^2(1-\varepsilon)^2-2\bar{\beta}A(1-\varepsilon)]}{2\bar{\beta}} > F$, or equivalently $A > \frac{\bar{\beta}(2-\varepsilon-\lambda)-\sqrt{\bar{\beta}[\bar{\beta}(3-2\varepsilon-\lambda)(1-\lambda)-2F(1-\varepsilon)^2]}}{(1-\varepsilon)^2}$, where $\frac{\bar{\beta}(2-\varepsilon-\lambda)-\sqrt{\bar{\beta}[\bar{\beta}(3-2\varepsilon-\lambda)(1-\lambda)-2F(1-\varepsilon)^2]}}{(1-\varepsilon)^2} < \overline{A^{Hp}}$ under Assumption $2.bis$. Then, the monopolist decides to

collect all patients' health data and develop project H when $A \geq \overline{A^{Hp}}$. From (13), we easily find that, in a corner solution, the firm offers $\delta(\beta) = \beta - A(1-\varepsilon)$ to patients with privacy types $\beta \in [\tilde{\beta}, \bar{\beta}]$, where $\tilde{\beta}(\bar{\delta}) = A(1-\varepsilon)$■

***Proof of Lemma* 1.**

It directly follows from comparing $d_p^* = \frac{\bar{\beta}A\varepsilon(1-\lambda)}{\bar{\beta}^2+(A)^2(1-\varepsilon)^2-2\bar{\beta}A(1-\varepsilon)(2-\lambda)}$ with $d^w = \frac{\varepsilon A}{\bar{\beta}-2A(1-\varepsilon)}$. ■

***Proof of Proposition* 5.**

The result follows from comparing d_p^* and d^w based on the ordering of private and social critical values of the incremental health benefit of developing project H. We hereby consider only cases where, under personalized compensation, the firm may overinvest in the new treatment (so that $0 < d_p^* \leq 1$ and $0 \leq d^w < 1$). Thus, we focus on the interval $A \in \left[\underline{A^{Hp}}, \overline{A^w}\right]$.

Consider first the case where $\underline{A^{Hp}} < \underline{A^w}$. For this to occur, we must have that the firm has strong bargaining power, that is, $0 < \lambda < \tilde{\lambda}(\varepsilon)$, where $\tilde{\lambda}(\varepsilon)$ is the smallest root of $\nu(\lambda) = (4 - 6\varepsilon + 2\varepsilon^2)\lambda^3 + (14\varepsilon - 4\varepsilon^2 - 11)\lambda^2 + (10 - 12\varepsilon + 4\varepsilon^2)\lambda - 3 + 4\varepsilon - \varepsilon^2 = 0$. We also must have that privacy costs are low enough, that is, $\bar{\beta} < \frac{2F(1-\varepsilon)^2\left(2-(2-\lambda)\lambda(1+2\lambda)\right)+4\sqrt{F^2(1-\varepsilon)^4(1-\lambda)^3(1-\lambda^2-2\lambda))}}{\varepsilon^2(2-\lambda)^2\lambda^2} \equiv b_2$.

In this case, when $\underline{A^{Hp}} < A < \underline{A^w}$, it easily follows that $d_p^* > 0 = d^w$. Then, when $A > \underline{A^w}$, we have to study two alternative orderings of critical values for A:

1. if $\underline{A^{Hp}} < \overline{A^{Hp}} < \underline{A^w} < \overline{A^w}$ then, when $\underline{A^w} < A < \overline{A^w}$, we have that $d_p^* = 1 > d^w > 0$.
2. if $\underline{A^{Hp}} < \underline{A^w} < \overline{A^{Hp}} < \overline{A^w}$ then, when $\underline{A^w} < A < \overline{A^{Hp}}$, we need to compare $d_p^* = \frac{\hat{\delta}_p^*}{\bar{\delta}}$ and $d^w = \frac{\delta^w}{\bar{\delta}}$. From Lemma 1, $\frac{\hat{\delta}_p^*}{\bar{\delta}} > \frac{\delta^w}{\bar{\delta}}$ when $\underline{A} < A < \overline{A}$. We find that, under ordering 2., $\underline{A} < \underline{A^w}$ and $\overline{A} > \overline{A^{Hp}}$. It follows that $\frac{\hat{\delta}_p^*}{\bar{\delta}} > \frac{\delta^w}{\bar{\delta}}$ whenever $\underline{A^w} < A < \overline{A^{Hp}}$. Furthermore, when $\overline{A^{Hp}} < A < \overline{A^w}$, we easily obtain that $d_p^* = 1 > d^w > 0$. Therefore, we can conclude that $d_p^* > d^w$ holds as long as $\underline{A^w} < A < \overline{A^w}$.

Now, consider the case where $\underline{A^{Hp}} > \underline{A^{w}}$. In such a case, there is only one ordering of critical values for A under which the firm may overinvest in data collection and thus in project H, namely, $\underline{A^{w}} < \underline{A^{Hp}} < \overline{A^{Hp}} < \overline{A^{w}}$. For this to occur, we must have that $0 < \lambda < \tilde{\lambda}(\varepsilon)$ and $\bar{\beta} > b_2$, or alternatively $\tilde{\lambda}(\varepsilon) < \lambda < \bar{\lambda} = \frac{3-4\varepsilon+\varepsilon^2}{4-4\varepsilon+\varepsilon^2}$. We find that, under this ordering of critical values, $\underline{A}$ is lower than $\overline{A^{Hp}}$ and it can be higher or lower than $\underline{A^{Hp}}$, whereas $\overline{\alpha} > \overline{\alpha^{w}}$. Hence, $\frac{\hat{\delta}_p^{\,*}}{\bar{\delta}} > \frac{\delta^{w}}{\bar{\delta}}$ holds when $max\{\underline{A^{Hp}}, \underline{A}\} < A < \overline{A^{Hp}}$. Moreover, when $\overline{A^{Hp}} < A < \overline{A^{w}}$, we easily find that $d_p^* = 1 > d^{w} > 0$.

Thus, to summarize the results obtained in the considered cases, the firm overinvests in treatment H as long as $0 < \lambda < \bar{\lambda}$ and $A \in I \subseteq \left[\underline{A^{Hp}}, \overline{A^{w}}\right]$. ■

***Proof of Proposition* 6.**

We compare private and social outcomes by sorting the relevant critical values of the incremental health benefit of the new treatment to develop project H, depending on other parameters. We find that, if $\lambda > \frac{\bar{\beta}-F(2-\varepsilon)}{\bar{\beta}} \equiv \lambda_2$ then $\overline{A^{w}} < \frac{F}{(1-\lambda)}$. Hence, there is no overinvestment in treatment H because, when private investment occurs, it is socially optimal to invest by using all patients' data. Instead, underinvestment arises when $A \in \left(\underline{A^{w}}, \frac{F}{(1-\lambda)}\right)$.

It follows that overinvestment may arise only if $\lambda < \lambda_2$. We distinguish the following cases:

(i) if $0 < 1 - \frac{F(1-\varepsilon)}{\bar{\beta}} - \frac{\sqrt{F[2F(1-\varepsilon)^2+\bar{\beta}\,\varepsilon^2]}}{\bar{\beta}\sqrt{2}} < \lambda < \lambda_2$, then we have $\underline{A^{w}} < \frac{F}{(1-\lambda)} < \overline{A^{w}}$. Thus, underinvestment arises when $A \in \left(\underline{A^{w}}, \frac{F}{(1-\lambda)}\right)$, whereas there is overinvestment when $A \in \left(\frac{F}{(1-\lambda)}, \overline{A^{w}}\right)$ because the firm invests by using too much data as compared to the social optimum (where $d^{w} < 1$). Otherwise, private incentives and social goals are aligned.

(ii) if $0 < \lambda < 1 - \frac{F(1-\varepsilon)}{\bar{\beta}} - \frac{\sqrt{F[2F(1-\varepsilon)^2+\bar{\beta}\,\varepsilon^2]}}{\bar{\beta}\sqrt{2}}$, then we have $\frac{F}{(1-\lambda)} < \underline{A^{w}} < \overline{A^{w}}$. Thus, overinvestment arises when $A \in \left(\frac{F}{(1-\lambda)}, \overline{A^{w}}\right)$. Specifically, when $A \in \left(\frac{F}{(1-\lambda)}, \underline{A^{w}}\right)$ the firm develops project H when

it is not socially optimal, whereas when $A \in \left(\underline{A^W}, \overline{A^W}\right)$ the firm invests by using too much data as compared to the social optimum. Otherwise, private incentives and social goals are aligned.

To conclude, there is overinvestment in treatment H when both $\lambda < \lambda_2$ and $\frac{F}{(1-\lambda)} < A < \overline{A^W}$ hold. Since λ_2 is increasing in $\bar{\beta}$ then overinvestment is more likely when $\bar{\beta}$ is high. ■

***Proof of Proposition* 7.**

Let us analyze separately the two cases where private incentives to invest in project H are misaligned under the policy and in the baseline model. In both cases, the firm invests by using all patients' data whenever they are freely available.

First, let $\frac{F}{(1-\lambda)} < A < \underline{A^H}$. Then, there is no investment in the baseline model. Hence, the policy improves social welfare if and only if $E(W_H|d=1) = A - \frac{\bar{\beta}}{2} + q_0 - F > q_0 = W_0$, that is, $A > \frac{\bar{\beta}}{2} + F$.

Note that, if $\bar{\beta} < \beta_l = \frac{2F\lambda}{1-\lambda}$, then we have $\frac{\bar{\beta}}{2} + F < \frac{F}{(1-\lambda)}$ and thereby social welfare is always higher under the considered policy when $A \in \left(\frac{F}{(1-\lambda)}, \underline{A^H}\right)$. Instead, if $\bar{\beta} > \beta_h = \frac{2F\left(\varepsilon\left(2(2-\lambda)-\varepsilon(1-\lambda)^2\right)+2\lambda\right)+4F\sqrt{2\varepsilon^2+4\varepsilon\lambda+\lambda^2-\varepsilon(2+\varepsilon)\lambda^2}}{\varepsilon^2(1-\lambda)^2} > \beta_l$ then we have $\frac{\bar{\beta}}{2} + F > \underline{A^H}$ and thereby social welfare is always lower under the policy when $A \in \left(\frac{F}{(1-\lambda)}, \underline{A^H}\right)$. Finally, if $\beta_l < \bar{\beta} < \beta_h$ then social welfare is higher (respectively, lower) under the policy when $A \in \left(\frac{\bar{\beta}}{2} + F, \underline{A^H}\right)$ $\left(A \in \left(\frac{F}{(1-\lambda)}, \frac{\bar{\beta}}{2} + F\right)\right)$.

Now, let $\underline{A^H} < A < \overline{A^H}$. Then, in the baseline model, the firm collects an amount of patients' data equal to $d^* = \frac{\varepsilon(1-\lambda)A}{2\left(\bar{\beta} - A(2-\lambda)(1-\varepsilon)\right)}$ (see Proposition 1). Hence, the policy improves social welfare if and only if $E(W_H|d=1) > E(W_H|d^*)$, that is $A - \frac{\bar{\beta}}{2} > \frac{A^2\varepsilon^2(1-\lambda)\left(\bar{\beta}(3+\lambda) - 2A(1-\varepsilon)(3-\lambda)\right)}{8\left(\bar{\beta} - A(1-\varepsilon)(2-\lambda)\right)^2}$, or, $A > \frac{\bar{\beta}(8-5\varepsilon-2\lambda+3\varepsilon\lambda) - \bar{\beta}\sqrt{9\varepsilon^2+2(10-7\varepsilon)\varepsilon\lambda+(2-3\varepsilon)^2\lambda^2}}{4(1-\varepsilon)(4-\varepsilon(1-\lambda)-2\lambda)} \equiv \tilde{A}$.

Since $\tilde{A} < \overline{A^H}$, then social welfare is higher under the policy at issue when the condition $max\{\underline{A^H}, \tilde{A}\} < A < \overline{A^H}$ holds. This means that, if $\tilde{A} < \underline{A^H}$ then when $A \in (\underline{A^H}, \overline{A^H})$ the policy always improves welfare. We find that this condition holds when $\bar{\beta} < \tilde{\beta}$, with $\tilde{\beta} =$

$\frac{F\left(8\lambda+3\varepsilon^2(1-\lambda^2)+2\varepsilon(6-\lambda-\lambda^2)\right)+F\sqrt{(4+\varepsilon(1-\lambda))^2(9\varepsilon^2+2(10-7\varepsilon)\varepsilon\lambda+(2-3\varepsilon)^2\lambda^2)}}{2\varepsilon^2(1-\lambda)^2}$. Finally, if $\bar{\beta} < \beta_l < \tilde{\beta}$ then the considered policy never reduces welfare. ■

***Proof of Proposition* 8.**

Let us first exclude the cases where investment in treatment H is totally precluded. This occurs when developing project H is not profitable even if the firm is granted free access to all patients' data. This means that $\Pi_H(d=1) < \Pi_0$ holds, that is, $A < \frac{F}{(1-\lambda)}$. Alternatively, this occurs when there is no feasible value of d such that $E(W_H) \geq W_0$, that is, when $A < \underline{A^w}$. In such a case, the firm is not allowed to access health data.

When $d^w \geq d^{min}$ then the agency's problem is equivalent to the welfare maximization problem in Section 5. Thus, the agency grants the firm free access to $d^w = \frac{\delta^w}{\bar{\delta}} = min\left\{\frac{\varepsilon A}{(\bar{\beta}-2A(1-\varepsilon))}, 1\right\}$, provided that $A \geq \underline{A^w}$.

Consider an interior solution of the welfare maximization problem, with $d^w < 1$ and $\underline{A^w} \leq A < \overline{A^w}$. Condition $d^w \geq d^{min}$ can be written as:

$$\Gamma \equiv -(A)^3\varepsilon^2(1-\varepsilon)(1-\lambda) + (A)^2\left(\bar{\beta}\varepsilon^2(1-\lambda) - 4F(1-\varepsilon)^2\right) + (A)4\bar{\beta}F(1-\varepsilon) - \bar{\beta}^2F \geq 0$$

Since $\frac{\partial\Gamma}{\partial A} > 0$ in the interval $[\underline{A^w}, \overline{A^w}]$, then the sign of Γ can be determined by studying the sign of the function evaluated at the boundaries of the interval. We thus find that:

i) if $0 < \lambda \leq \frac{1}{2} + \sqrt{\frac{F(1-\varepsilon)^2}{4F(1-\varepsilon)^2+2\bar{\beta}\varepsilon^2}} \equiv \lambda_1$, then $(\Gamma|A = \underline{A^w}) \geq 0$, so that $d^w \geq d^{min}$. Since $\frac{F}{(1-\lambda)} < \underline{A^w}$, the first best is achieved in the whole interval $[\underline{A^w}, \overline{A^w}]$.

ii) if $\lambda_1 < \lambda < \frac{\bar{\beta}-F(2-\varepsilon)}{\bar{\beta}} \equiv \lambda_2$, then $\left(\Gamma\middle|A = \underline{A^w}\right) < 0$ and $\left(\Gamma\middle|A = \overline{A^w}\right) > 0$. Thus, there is a critical value $\hat{A} \in (\underline{A^w}, \overline{A^w})$ such that $d^w = d^{min}$ and the first best is achieved when $\hat{A} \leq A \leq \overline{A^w}$. Since $\frac{F}{(1-\lambda)}$ can be higher or lower than $\underline{A^w}$, then $d = d^{min}$ might solve the agency's problem when $max\left\{\underline{A^w}, \frac{F}{(1-\lambda)}\right\} \leq A \leq \hat{A}$.

iii) if $\lambda \geq \lambda_2$, then $\left(\Gamma\middle|A = \overline{A^w}\right) \leq 0$ and $d^w < d^{min}$ in the whole interval $[\underline{A^w}, \overline{A^w}]$. Nonetheless, since $\frac{F}{(1-\lambda)} > \overline{A^w}$ then $d = d^{min}$ cannot be a solution of the agency's problem.

Consider now a corner solution where $d^w = 1$ and $A > \overline{A^w}$. We find that, if $\lambda < \lambda_2$ then $\frac{F}{(1-\lambda)} < \overline{A^w}$, so that $d^w > d^{min}$ and the first best is achieved when $A > \overline{A^w}$. Instead, if $\lambda \geq \lambda_2$ then $\frac{F}{(1-\lambda)} > \overline{A^w}$, so that $d^w > d^{min}$ and the first best is achieved when $A > \frac{F}{(1-\lambda)}$. Otherwise, the policy is not applied because it does not improve welfare.

Now, let $\lambda_1 < \lambda < \frac{\bar{\beta}-F(2-\varepsilon)}{\bar{\beta}} \equiv \lambda_2$ and $max\left\{\underline{A^w}, \frac{F}{(1-\lambda)}\right\} \leq A \leq \hat{A} < \overline{A^w}$. Thus, d^{min} does solve the agency's problem when $E(W_H|d = d^{min}) \geq W_0 = q_0$. Condition $E(W_H|d = d^{min}) \geq W_0$ can be written as:

$$\mathrm{Z} \equiv \frac{\bar{\beta}\varepsilon}{4(1-\varepsilon)^2}\left(\sqrt{\varepsilon^2 + \frac{4F(1-\varepsilon)}{A(1-\lambda)}} - \varepsilon\right) - \frac{F\left(\bar{\beta} - 2A(1-\varepsilon)\lambda\right)}{2A(1-\varepsilon)(1-\lambda)} > 0$$

with Z increasing in A as long as $max\left\{\underline{A^w}, \frac{F}{(1-\lambda)}\right\} \leq A \leq \overline{A^w}$. As above, we study the sign of Z evaluated at the boundaries of the interval and find that the sign is negative (respectively, positive) at the lower (upper) bound. Then, it does exist a critical value $\bar{\bar{A}} \in [max\left\{\underline{A^w}, \frac{F}{(1-\lambda)}\right\}, \overline{A^w}]$ such that $\mathrm{Z}(\bar{\bar{A}}) = 0$. We find that $\bar{\bar{A}} < \hat{A}$. Indeed, if $A = \hat{A}$ then $d^w = d^{min}$. Hence, $\mathrm{Z}(\hat{A}) > 0$ because $A > \underline{A^w}$ ensures that $E(W_H|d = d^w) > W_0$. It follows that, when $\bar{\bar{A}} \leq A \leq \hat{A}$, social welfare improves relative to the *status quo* when the firm invests in treatment H by using a larger amount of data than the socially optimal amount. Instead, when $max\left\{\underline{A^w}, \frac{F}{(1-\lambda)}\right\} < A < \bar{\bar{A}}$ the agency grants no access to data and there is no investment.

It is straightforward to show that, under the proposed policy, social welfare is at least as high as in the baseline model. Indeed, the condition $E(W_H) > W_0$ ensures higher welfare when there is investment under the policy and no investment in the baseline model. When instead the firm invests by using the suboptimal amount of data d^* in the baseline model, the policy achieves first best because $d^{min} < d^* < d^w$ holds. ■