Regulation of Selection Technologies*

Marie Obidzinski[†] Marianne Verdier[‡]

May 7, 2026

Abstract

We analyze how a monopoly chooses the quality of a selection technology designed to exclude non-compliant consumers, who may be more costly to serve. The firm may decide to exclude a consumer after observing a signal received on her compliance, the accuracy of which depends on the quality of the technology. The choice of the quality of the selection technology also affects the incentives of consumers to comply with legal rules and buy a higher-quality service. We show that the monopoly distorts the choice of the quality of the selection technology, which results in an inefficient participation of consumers in the higher-quality market and an inefficient allocation of consumers across the compliant and non-compliant categories, respectively. We analyze the role of sanctions for mis-classification errors to improve the allocation. We show that a sanction for false negatives may sometimes reduce the surplus of compliant consumers. We compare the effect of several complementary instruments (e.g. price cap, fine on non-compliant consumers, sanction for false positive) on welfare.

Keywords: compliance, selection markets, monetary sanctions, regulation of algorithms.

JEL classification: K4, L51, 031.

*We thank Christophe Hurlin, Laurent Linnemer, Maarten van Oordt, Elena Dumitrescu, the participants to the 40th symposium on Money, Banking and Finance of the GDRE conference (July, 2024), to the 16th Paris Conference on Digital Economics, to the 42nd annual Conference of the European Association of Law & Economics (September, 2025), to the 10th annual Conference of the French Association of Law and Economics (October, 2025), and to the Risk Forum at Institut Louis Bachelier (March, 2025) for their insightful comments on the paper. All remaining errors are ours.

[†]Paris-Panthéon-Assas University, CRED UR 7321, 75005 Paris, France. e-mail: marie.obidzinski@assas-universite.fr

[‡]Paris-Panthéon-Assas University, CRED UR 7321, 75005 Paris, France. e-mail: marianne.verdier@assas-universite.fr

1 Introduction

The rise of artificial intelligence (AI) gives firms the ability to use advanced statistical techniques to collect and process consumer data. In several sectors of the economy, this technology helps firms define consumer categories, sort them, and sometimes even exclude them from a market. The use of a *selection technology* to exclude consumers is particularly widespread in markets where it is necessary to comply with legal rules or standards to buy a product. For example, it is necessary to perceive revenues from legal activities to open a bank account, or firms may sometimes need to comply with environmental standards to receive credit.

However, the consumer selection process is associated with many costly errors. Some consumers who are compliant may be forced to renounce purchasing a service when they are excluded by a type I error (false positive), whereas non-compliant ones may create social damage when they are included by a type II error (false negative). Errors generate externalities across consumer categories. In particular, an algorithm can efficiently detect non-compliant consumers, but this creates negative externalities on compliant ones, who are more often excluded. In addition, errors can indirectly affect the volume of consumers who belong to each classification category. They could even have the perverse effect of reducing consumer incentives to comply, because their effort is not rewarded by the possibility of purchasing their preferred service.

Firms control the volume of errors by choosing the performance of their selection technology. However, they make this decision according to the profit they make from each category of consumers. Sometimes, non-compliant consumers have a higher willingness to pay for their services than compliant ones, so it is not profitable to exclude them completely, even though they cause social damage. Thus, the widespread adoption of complex selection technologies can have substantial welfare effects that deserve to be better understood.

In this paper, we show that a monopoly chooses an inefficient performance for the selection technology, which results in a distortion of the total volume of sales and their allocation between compliant and non-compliant consumers, respectively. We offer an economic measure of the welfare effects of the monopoly's choice of a selection technology that takes into account the consumers' responses, both in terms of incentives to comply and participation in the market. We analyze how sanctions for mis-classification errors impact the monopoly's incentives to select by quantity and technology.

Firms often use classification algorithms in selection markets when they incur higher costs of selling to some consumer types, which are costly to observe. If a firm uses an algorithm, it can still exclude consumers who are less profitable by changing prices (Akerlof, 1970, Einav and Finkelstein, 2011). Veiga and Weyl (2016) label this effect as *selection by quantity*. However, firms also rely on the quality of the selection technology to expand its ability to sort profitable consumers from non-profitable ones. The quality of the prediction delivered by the selection technology changes the consumers' marginal willingness-to-pay for the product for a given price according to their decision to comply with the law. This is because sophisticated consumers may anticipate that, with some probability, the firm will prevent them from purchasing its product. Therefore, firms can also rely on *sorting by technology*.

We motivate our model by the example of fraud detection algorithms and Anti Money Laundering (AML) policies that are widely adopted in several sectors of the economy, such as banking and finance.[1] Banks have incentives to screen their consumers because they risk paying sanctions if they serve criminals. This implies that they incur different costs of serving honest consumers and criminals, respectively. In this context, a better-quality screening technology deters criminals from engaging in illegal activities, but is also associated with the inefficient exclusion of compliant consumers.[2]

We model a market in which consumers benefit from not complying with the law or a standard. However, compliance is required to purchase a higher-quality product or service from a seller. The monopolistic seller offers a service at a uniform price and does not observe whether a consumer is

[1]The model could also apply to credit markets because firms can make an effort to try and belong to a low-risk category if this increases their probability to obtain some funding. A better-quality screening technology may give entrepreneurs an incentive to make an effort to respect a standard (e.g, ESG) before submitting their credit application.

[2]For example, van Oordt (2025) mentions that not-for-profit organizations are facing challenges in opening or maintaining bank accounts in various jurisdictions because banks prefer to de-risk and adopt strict exclusion rules (e.g., Commission, 2024; Dutch Association for Charities, 2024).

compliant or not. Before selling its services, it relies on a screening technology that delivers a signal on the consumer's compliance. The signal is imperfectly informative, but enables a classification of consumers into two categories: compliant and non-compliant, respectively. Being aware of this choice, the consumers may decide to become either compliant or non-compliant. Therefore, the selection technology affects the segmentation of consumer types in the market, which the firm does not observe. Non-compliant consumers may sometimes have a higher willingness-to-pay for the firm's services, which creates adverse selection. The firm may select its consumers by increasing its price or the quality of the selection technology. Consumers who decide to become non-compliant generate an externality on compliant ones because they impact the price that the latter pay for the service and their probability of being excluded.

We show that the monopoly distorts the choice of the quality of the selection technology with respect to the social optimum, which results in the offering of an inefficient total quantity to consumers, and an inefficient allocation of consumers between the compliant and non-compliant categories, respectively. This implies that sometimes the monopoly does not sell to the socially optimal set of consumer categories. For example, the monopoly may choose to de-risk and completely exclude non-compliant consumers, though it would be socially optimal to sell to both categories of consumers. Or, even if the monopoly sells to both consumer categories when it is socially optimal of doing so, it excludes an inefficient relative share of compliant consumers with respect to non-compliant ones. This is because it only takes into account the impact of the marginal variation of the exclusion probability on its cost, without perfectly internalizing the social damage and the externalities across consumer categories.

We compare our model with the literature on platform markets, in which an intermediary is able to choose different prices for categories of users who exert network externalities on each other. A monopolistic platform does not choose the socially optimal structure of prices, which generates an inefficient relative participation of each group of consumers (Rochet and Tirole, 2006, Weyl, 2010). In a selection market, it is not possible to observe the categories of users ex ante and choose different prices for each of them. However, a monopoly can also affect the relative participation of each group of users by choosing the quality of the selection technology and the price of its service. This decision creates indirect participation externalities between compliant and non-compliant consumers, which implies that the monopoly behaves as a platform choosing how much to sell to its consumers and how to allocate its sales between compliant and non-compliant consumers, respectively.[3]

We use our framework to offer a measure of the welfare effects of fraud detection algorithms, which differs from the existing literature. We show that the usual evaluations of classification algorithms generally neglect how the choice of a performance for the algorithm endogenously affects the participation of each consumer category in a market and their consumption surplus.

We complete our analysis by analyzing the role of sanctions for mis-classification errors. Supervisors often rely on sanctions for type II errors, when a firm erroneously sells a service to non-compliant consumers, but more rarely on sanctions for type I errors, when a firm excludes a compliant consumer by error. We show that if the price of the service and the quality of the selection technology are substitute strategies to increase the monopoly's profit, higher sanctions for type II errors increase the monopoly's price but reduce the quality of the selection technology. However, if the quality of the selection technology and the price are complements, higher sanctions for type II errors may have unintended effects and even sometimes reduce the monopoly's price. Due to externalities across consumer categories, the regulator needs at least two instruments to incentivize the monopoly to sell to the socially optimal set of consumer categories. We compare several instruments: a price cap, a sanction for type I errors, and a fine imposed on non-compliant consumers. We conclude the analysis by showing that a sanction for type II errors sometimes has the perverse effect of reducing the surplus of compliant consumers.

The remainder of the paper is organized as follows. Section 2 surveys the literature in connection with our study. Section 3 develops the model. Section 4 determines the firm's choice of the quality of the selection technology. Section 5 compares the firm's choice with the first-best and analyzes the role of sanctions for errors. Section 6 concludes.

[3]This comparison makes a connection between the literature on platform markets (Weyl, 2010) and the literature on contracting with externalities (Segal, 1999).

2 Literature

Our paper offers a model for analyzing the regulation of a monopoly, which undertakes a selection activity. The quality of the selection technology impacts the consumers' expected utility of buying a higher-quality service rather than an outside option. Therefore, our framework shares similarities with the model of Spence (1975), which analyzes the regulation of the quality offered by a monopoly.[4] In contrast, we consider that the quality of the product is exogenous and analyze the endogenous choice of the quality of a selection technology. The monopoly's selection activity impacts consumer ability to buy the service, her compliance decision and indirectly generates social damage. In this context, the regulation aims to give the monopoly incentives to select compliant buyers, who generate lower social damage. Colliard (2019) analyzes the optimal regulation of a monopoly that can hide a selection model in the credit market. We consider instead that the choice of the selection model is endogenous and impacts the segmentation of consumer types. Besfamille et al. (2025) consider the optimal regulation of a monopoly that makes an effort to detect evaders who do not pay the price for its service. The monopoly is unable to exclude evaders ex ante. In contrast, we consider that the monopoly screens its consumers before allowing them to buy its service and may exclude them.

The literature on price discrimination analyzes how firms can segment markets by offering menus of contracts with different prices and qualities (Stiglitz, 1977, Mussa and Rosen, 1978). In selection markets, firms also use uniform pricing with non-price features such as product quality or a collateral to select profitable consumers (Einav et al., 2010, Einav and Finkelstein, 2011, Mahoney and Weyl, 2017). We adopt the same approach and extend, in particular, the model of Veiga and Weyl (2016) by assuming that a firm chooses the quality of a selection technology and a uniform price for a service. We also consider that the firm incurs different costs when serving heterogeneous types of consumers and needs to attract the right users to be profitable (see Veiga et al. (2017) or Biancini and Verdier (2023)).

[4] See Besanko et al. (1987) for an analysis of the regulation of a monopoly, when consumers differ in their willingness to pay for quality. In our paper, the distribution of consumer types is endogenous and depends on the choice of the quality of the selection technology.

The issue of consumer selection is related to the broader analysis of the increasing role of private firms as gatekeepers.[5] In several sectors, public authorities tend to delegate law enforcement to private firms. The focus is not on the harm that the private firm directly inflicts on society, but on the harm caused by users, whom Spier and Van Loo (2025) call "bad actors". However, financial intermediaries or technology platforms can influence the proportion of bad actors by investing in a selection technology. The recent literature on platform liability (Creti and Verdier, 2014, Hua and Spier, 2025, 2023) has focused on the allocation of damage between the users and the platform, while we focus on the impact of the sanctions for mis-classification errors on consumer selection.

Selection costs can be related to a literature that incorporates data as input in the firms' production function (see Farboodi and Veldkamp, 2020). The investment in quality of a selection technology could be interpreted as the choice to collect more consumer data (as in Gurkan and de Véricourt, 2022), while the probability of being excluded for the consumer causes an inconvenience cost of data collection (see, for instance, Markovich and Yehezkel (2021)).[6] We contribute to this literature by considering that the inconvenience cost of data collection is group-specific and depends on the firm's choice of a selection technology. Several articles consider that strategic consumers exert externalities on each other when they decide to share their data (Garratt and Van Oordt, 2021, Acemoglu et al., 2022).[7] In our framework, there are also externalities across consumer categories because a consumer's decision to become non-compliant impacts the price paid by compliant ones and their probability of being excluded.

Artificial Intelligence (AI) is an example of a technology that can be used to select consumers more efficiently than human analysis (see Cowgill et al., 2025, Goh and Lee, 2019).[8] However, the usual measures of algorithmic performance such as accuracy fail to account that the quality of the

[5]See, for instance Van Loo (2020), which highlights the role of large firms as new gatekeepers and in particular their role *vis-à-vis* third parties.

[6]In the literature, the firm's incentives to collect data may depend on the price discrimination possibilities (Bergemann and Bonatti, 2019, Ichihashi, 2021), the individuals' decision to share their personal data (O'Brien and Smith, 2014, Jullien et al., 2020, Acemoglu et al., 2022), and competition (Jones and Tonetti, 2020).

[7]A literature in computer science analyzes how strategic consumers can manipulate the information given to a selection technology (i.e., an algorithm) so as to impact the outcome of the classification process, and whether it is possible to design learning processes that are robust to potential data manipulation (See Dong, et al., Hardt et al.).

[8]An important aspect of AI adoption is related to the interactions between technology and human judgment in decision-making which is not the focus of our paper (Agrawal et al., 2018, Daugherty and Wilson, 2018, Mullainathan and Spiess, 2017, Kleinberg et al., 2017).

screening technology impacts consumer demand.[9] Our paper contributes to a scarce theoretical literature that focuses on the interactions between AI adoption and product pricing. In a close work, Gans (2023) analyzes the adoption of AI by a monopoly facing demand uncertainty when it chooses its price and quantity. Gurkan and de Véricourt (2022) consider the interplay between data collection, algorithmic performance and product pricing in a two-period model. We consider instead how the simultaneous choice of the price for the service and the quality of the selection technology affects consumer incentives to comply and the welfare effects of the firm's errors.

The literature on AI regulation analyzes different remedies to correct inefficient adoption of AI, or reduce the damage caused by the technology. We contribute to this strand of the literature by studying how sanctions for mis-classification errors can influence the choice of the performance of an algorithm, product pricing, and consumer incentives to comply. Acemoglu and Lensman (2024) analyze how sector-independent taxes or sandboxes affect the dynamics of AI adoption in the presence of negative external effects. Rambachan et al. (2020) determine how a social planner should choose a fair algorithm, when a group of disadvantaged users exhibits observable characteristics that may imply their exclusion. We instead focus on the role of moral hazard, when the quality of the technology affects the consumers' incentives to comply and their probability of being excluded. Chen and Hua (2026) focus on how liability rules can enhance AI safety by encouraging investments in R&D.[10] Battiston et al. (2024) offer a method to rank different tax evasion detection algorithms with heterogeneous costs of type I and type II errors, but do not consider the interactions of AI regulation and product pricing.

[9] See Fraisse and Laporte (2021) and Hurlin et al. (2024) for examples of assessments of AI performance in credit scoring, or see Zhang and Trubey (2019) for anti-money laundering (AML).

[10] See also Llanes and Madio (2024) for an analysis of the effect of liability rules on AI adoption. In the related field of performative algorithms, Dawid and Muehlheusser (2022) highlight the trade-off of increasing producer liability between incentivizes him to invest more in safety, but reduces the algorithm market penetration (here, autonomous vehicules). Obidzinski and Oytana (2024) also take into account the impact of producer liability on the diffusion of AI technology.

3 The model

We build a model to analyze the regulation of a monopoly that may select its consumers by quantity and technology. The choice of the performance of the selection technology impacts the segmentation of consumer types and their demand for the monopoly's service.[11]

Firms and selection technology

A monopoly offers a service of exogenous quality $\Delta > 0$ at a price p, whereas a competitive fringe of firms sells a service of quality $\Delta = 0$ at their marginal cost which is equal to zero. Firms may sell their services to two different consumer categories indexed by $i \in \{c, nc\}$: compliant ($i = c$) or non-compliant ($i = nc$). Before selling their service, the monopoly may rely on a costly selection technology such as an algorithm to determine whether a consumer is non-compliant and exclude her from the market. Firms from the competitive fringe do not use any selection process and, therefore, do not exclude non-compliant consumers.[12]

The monopoly is able to parametrize the selection process by choosing a variable s in an interval $[0, S]$ of $\mathbb{R}_+$, which we will refer to as the quality of the selection technology.[13] If the monopoly excludes a consumer, she is forced to renounce buying the higher-quality service and instead consumes the lower-quality service from the competitive fringe of firms. Given the quality s of the selection technology, the monopoly excludes a consumer belonging to category $i \in \{c, nc\}$ with probability $e(s, \theta) \equiv e_i(s)$, and agrees to sell her the higher-quality service with probability $1 - e_i(s)$, where $0 \leq e(s, \theta) \leq 1$. With probability $e_c(s)$, the monopoly excludes a compliant consumer (a false positive) and makes a type I error, whereas with probability $1 - e_{nc}(s)$, it sells to a non-compliant consumer (a false negative) and makes a type II error.[14]

If $s = 0$, the monopoly has the same probability $e_0 \in [0, 1]$ of excluding a compliant consumer and a compliant consumer, respectively, and we have $e_c(0) = e_{nc}(0) = e_0$. If $s = S$, the monopoly

[11]As shown in Appendix A-3, our framework is an adaptation of the model of Veiga and Weyl (2016).

[12]Unlike in the literature on rational inattention, we abstract from analyzing the role of the firm's ex ante belief on the consumer's identity (see Sims, 2003, and Maćkowiak et al., 2023, for surveys).

[13]In Appendix G, we give common empirical measures of algorithmic performance.

[14]The null hypothesis corresponds to the assumption that a consumer is not compliant.

excludes a non-compliant consumer with certainty, that is, we have $e_{nc}(S) = 1$, and does not systematically exclude compliant consumers, that is, we have $e_c(S) < 1$.

We formalize two assumptions on the selection technology.

(A1) there exists $\overline{s} \in [0, S]$ such that a better-quality selection technology increases the probability of excluding a non-compliant consumer, that is, we have $\partial e_{nc}/\partial s \geq 0$.

(A2) A better-quality selection technology provides consumers with incentives to comply, that is, we have $\partial(e_{nc} - e_c)/\partial s \geq 0$.

A better-quality screening technology may have heterogeneous effects on the probabilities of excluding different types of consumers, and it is not necessarily monotonous.

Consumer choices

Following a large literature on crime deterrence (Becker, 1968), we consider a continuum of risk-neutral consumers, who differ with respect to their private benefit b of becoming non-compliant.[15] We assume that b is distributed on $[0, B]$ according to the probability density f, with cumulative distribution F, where $B > 0$.[16] Consumers have a willingness to pay for one unit of a service of quality $\Delta \geq 0$ given by $(1 + \Delta)(y + lb)$, where $y > 0$ is their legal revenue and $b > 0$ is their benefit of being non-compliant, with $l = 0$ if a consumer is compliant ($i = c$) and $l = 1$ if she is not compliant ($i = nc$). Consumers make two consecutive decisions:

- They decide whether or not to comply with the legal framework. After this decision, which is irreversible, the consumer is characterized by her category $i \in \{c, nc\}$ and her initial type b. Because the firm does not observe two pieces of information on the consumer, we will refer to $\theta = (b, i)$ as being the consumer's bi-dimensional type.

- They decide from which firm to buy the service given their observation of the quality s of the

[15]The benefit b could also represent the consumer's benefit of hiding some personal information that increases a firm's cost of data collection (see Markovich and Yehezkel (2021)).

[16]This is a classic hypothesis in the public law enforcement literature issued from Becker (1968). See the surveys by Garoupa (1997) and Polinsky and Shavell (2007).

selection technology.[17] Consumers anticipate that if they buy the service from the monopoly and belong to category $i \in \{c, nc\}$, they may be excluded by the firm's selection process with probability e_i and unable to purchase the higher-quality service.

In case the monopoly detects a non-compliant consumer (a true positive), it systematically reports it to the regulator, and the consumer has to pay a fine $\phi \geq 0$, which is collected by the regulator at no cost. In that case, the consumer's net benefit of being non-compliant and purchasing the outside option is $y + b - \phi$.[18] If the monopoly excludes a compliant consumer (a false positive), the compliant consumer is constrained to purchase the lower-quality service and gets y.[19] We formalize an additional assumption:

(A3) $e_0 \leq \frac{\Delta(B+y)}{\phi+\Delta(B+y)}$.[20]

Assumption (A3) ensures that the demand of non-compliant consumers is not systematically equal to zero. Given that $e_{nc}(S) = 1$ and $e_c(S) < 1$, the monopoly can sometimes exclude non-compliant consumers while serving compliant ones.

Indirect expected utilities of consumers in category $i \in \{c, nc\}$ who expect to purchase the monopoly's service are given by

$$\mathcal{U}^i(p, e_i, b) \equiv \tilde{u}(e_i, b, i) - (1 - e_i)p, \quad (1)$$

where the expected gross utility of the consumer to purchase the monopoly's service is given by

$$\tilde{u}(e_i, b, i) \equiv (1 - e_i)(1 + \Delta)(y + lb) + e_i(y + l(b - \phi)),$$

with $l = 1$ if the consumer is not compliant ($i = c$) and $l = 0$ otherwise ($i = nc$). Consumers who expect to buy the outside option do not need to comply because there is no selection process, and

[17] This assumption can easily be relaxed by assuming that only a proportion $\eta \in (0, 1)$ of consumers is sophisticated.

[18] Naturally, the regulator's ability to impose a fine on non-compliant consumers depends on whether the latter have deep pockets or are judgment-proof, and whether it is costly to enforce the legislation. These cases can be captured in our setting by making some comparative statics with respect to ϕ.

[19] We neglect here all other costs and benefits that could be obtained by a consumer when she is excluded.

[20] See Appendix B-3 for the details of this assumption.

their indirect expected utilities of purchasing the lower-quality service ($\Delta = 0$) are given by

$$\mathcal{U}^o(b) \equiv y + b. \tag{2}$$

Consumer demand

If the monopoly chooses a price and a selection technology such that the market is not covered, consumers trade off between buying the higher-quality service from the monopoly and the lower-quality service from the competitive fringe of firms. For $i \in \{c, nc\}$, we denote by $\theta_i = (b_i, i)$ the type of the consumer indifferent between these two options, and it is implicitly defined by $\mathcal{U}^i(p, e_i, b_i) = \mathcal{U}^o(b_i)$.

The set of marginal buyers of category $i \in \{c, nc\}$ is defined as $\mathcal{M}_i \equiv \{b : b = b_i\}$, and the set of marginal buyers is $\mathcal{M} \equiv \mathcal{M}_c \cup \mathcal{M}_{nc}$. We also denote by $\mathcal{B}_i$ the set of buyers in category $i \in \{c, nc\}$ who prefer to buy the service offered by the monopoly rather than the outside option if they are not excluded, with $\mathcal{B} \equiv \mathcal{B}_c \cup \mathcal{B}_{nc}$. Similarly, the set of consumers in category $i \in \{c, nc\}$ who prefer the outside option is denoted by $\mathcal{O}_i$. The monopoly expects to receive a demand

$$D_i(s, p) \equiv \int_{\mathcal{B}_i} f(b) db$$

from consumers of category $i \in \{c, nc\}$. Total consumer demand for the higher-quality service is

$$Q(s, p) \equiv D_c(s, p) + D_{nc}(s, p).$$

As we shall demonstrate later in the analysis, since Q is decreasing with p, there exists a differentiable inverse demand $P(s, q)$ such that $Q(s, P(s, q)) = q$ for any $s \in [0, S]$.

If the market is covered (i.e, $Q(s, p) = 1$), all consumers purchase the higher-quality service, but there is a marginal type of consumer b_I who is indifferent between being compliant and non-compliant and it is implicitly defined by $\mathcal{U}^c(p, e_c, b_I) = \mathcal{U}^{nc}(p, e_{nc}, b_I)$.

The costs of the monopoly

If it serves a consumer in category $i \in \{c, nc\}$, the monopoly incurs a marginal cost $k(s, \theta) \equiv k_i(s)$, while if it excludes a consumer in category $i \in \{c, nc\}$ and misses a sale, the monopoly incurs a marginal cost $m(s, \theta) \equiv m_i(s)$. These marginal costs are the sum of:

1. **the marginal cost of data collection:** $a(s, \theta) \equiv a_i(s)$ to process the selection algorithm, which is spent whether the monopoly serves a consumer in category $i \in \{c, nc\}$ or whether it excludes her.[21] Improving the selection process involves collecting more data on each consumer. Therefore, we assume that the marginal cost of collecting consumer data increases with the quality of the selection technology; that is, we have $\partial a_i/\partial s \geq 0$, with $a_i(0) = 0$.

2. **the marginal costs of errors:** given by $r(\theta) \equiv r_i \geq 0$ for consumers who are not excluded and by $t(\theta) \equiv t_i \geq 0$ for consumers who are excluded, respectively, for $i \in \{c, nc\}$. We assume that these costs only depend on the regulatory sanctions for type I and type II errors, respectively, with $r_c = 0$ and $t_{nc} = 0$.[22]

Therefore, for each category $i \in \{c, nc\}$, we have $k_i(s) = a_i(s) + r_i$ and $m_i(s) = a_i(s) + t_i$, with $r_c = 0$ and $t_{nc} = 0$.

The firm's objective

If the monopoly chooses the selection technology, we assume that its objective consists of maximizing its profit with respect to s and q (or s and p), that is,

$$\tilde{\pi} \equiv \sum_{i \in \{c,nc\}} \pi(\theta, s, q), \tag{3}$$

with $\pi(\theta, s, q) = \pi^i(s, q) \equiv \int_{\mathcal{B}_i} \big((1 - e_i(s))(P(s, q) - k_i(s)) - e_i m_i(s)\big) f(b) db$.

[21] This cost is group-specific, because some information may be more difficult to obtain from non-compliant consumers, or because data processing costs cannot be symmetrically allocated to each category of consumers.

[22] We assume that it is impossible to subsidize the participation of either category of consumers. We also abstract from reputation effects that could increase the firm's marginal cost of selling to non-compliant consumers, which are standard and could be easily added to our model without changing the effects discussed in our paper.

The social planner's objective

A non-compliant individual causes social damage of value $d_i(b)$ if she consumes the higher-quality service and $d_0(b)$ if she consumes the outside option, respectively. The gross contribution to welfare of an individual of type $\theta = (b, i)$ is the sum of her gross utility of consuming the service less the cost of social damage, and it is given by

$$w(\theta) \equiv w_i(b) = (1 + \Delta)(y + l(b - d_i(b))),$$

if she consumes the higher-quality service without being excluded and

$$o(\theta) \equiv o_i(b) = y + l(b - d_o(b)),$$

if she consumes the outside option, respectively, where $l = 1$ if she is non-compliant, $l = 0$ otherwise.[23] Total social welfare given the monopoly's response is

$$W(s, q) = \sum_{i \in \{c, nc\}} \int_{\mathcal{B}_i} \tilde{w}_i(s, b) f(b) db + \mathbb{E}(o(\theta)), \tag{4}$$

where

$$\tilde{w}_i(s, b) \equiv (1 - e_i(s))(w_i(b) - \lambda r_i - o_i(b)) - e_i(s)\lambda t_i - a_i(s). \tag{5}$$

In Eq.(5), we assume that, due to detection and audit costs, the regulator only collects a share $1 - \lambda$ of the revenues from the sanctions for mis-classification errors, with $\lambda \in (0, 1)$. We neglect the cost of collecting fines from the non-compliant consumers.

To study the most interesting case, we make the following assumption:

(A4) there exists $s \in [0, S]$ and $p \in [0, \Delta y]$ such that the society may benefit from having a

[23] Our approach allows for different levels of harm per consumers having different benefits from non compliance. In the literature issued from Becker (1968), the level of harm induced per illegal act is usually assumed to be identical.

higher-quality service with a selection technology, that is, we have $W(s,q) > \underline{W}$, where

$$\underline{W} \equiv \int_0^B o_{nc}(b) f(b) db \tag{6}$$

represents social welfare if all consumers are non-compliant.

Additional notations

To make our results comparable to their paper, we follow the notations used by Veiga and Weyl (2016). If the market is not covered, we define the marginal impact of a price increase on the set of marginal buyers in category $i \in \{c, nc\}$ by

$$M_i(s,p) = (-1)^{l+1} \frac{\partial b_i}{\partial p} f(b_i) > 0, \tag{7}$$

where $l = 1$ if $i = nc$ and $l = 0$ if $i = c$.[24] Then we have $M(s,p) = -\partial Q/\partial p = M_c(s,p) + M_{nc}(s,p)$. The marginal consumer surplus is

$$MS \equiv \frac{Q}{M}.$$

We define two additional expectation operators that will be useful for our analysis. For an arbitrary smooth function $z(s,\theta)$, the expectation conditional on the set of buyers $\mathcal{B}$ is given by:

$$\mathbb{E}\left[z(s,\theta)\,|\mathcal{B}\right] = \frac{1}{Q} \sum_{i \in \{c,nc\}} \int_{\mathcal{B}_i} z(s,\theta) f(b) db. \tag{8}$$

We also define the expectation of any $z(s,\theta)$ conditional on the set of marginal consumers:

$$\mathbb{E}\left[z(s,\theta)\,|\mathcal{M}\right] \equiv \frac{1}{M} \sum_{i \in \{c,nc\}} (-1)^{(l+1)} \frac{z(s,b_i,i)}{\frac{\partial \tilde{u}(s,b_i,i)}{\partial b} - 1} f(b_i), \tag{9}$$

[24] We use opposite signs because we will see that b_c decreases with the price p, whereas b_{nc} increases with the price p (see later Eq.(10)).

where $l = 1$ if $i = nc$ and $l = 0$ if $i = c$. We use Newton's notation to denote partial derivatives with respect to s, namely

$$z'(s,\theta) = \frac{\partial z(s,\theta)}{\partial s}.$$

Timing of the game

The timing of the game is as follows: First, nature chooses the magnitude of the private non-compliance benefit b obtained by a consumer. Second, the monopoly (or the regulator) chooses the quality of the selection technology s and the price p for the higher-quality service. Third, each individual learns his private benefit of being non-compliant b and after observing p and s, decides whether or not to comply with the existing legal framework, given her expectations on the possibility of buying the higher-quality service. A consumer in category $i \in \{c, nc\}$ expects that she will obtain $\mathcal{U}^i(p, e_i, b)$ if she purchases the higher-quality service offered by the monopoly and $\mathcal{U}^o(b)$ if she purchases the outside option, respectively. Fourth, consumers who demand the monopoly's service are subject to a screening process. Those consumers who are not excluded benefit from the higher-quality service, whereas the others purchase the outside option. Fifth, the regulator audits the firm's selection process and may impose sanctions for mis-classification errors. The judge punishes the non-compliant consumers if the latter have been detected by the firm.

In the next section, we solve the game through backward induction.

4 Profit maximization

In this section, we determine the firm's private choice of the quality of its selection technology and the price of its service.

4.1 Consumer demand and the selection strategies

We first derive the consumer decision to be compliant and her demand for the monopoly's service given her expectation of the risk of being excluded from the higher-quality market. If the higher-quality market is not covered, a consumer trades off between buying the higher-quality service from the monopoly and purchasing the lower-quality outside option. From (1) and (2), if $e_i < 1$, the type of consumer in category $i \in \{c, nc\}$ who is indifferent between purchasing the monopoly's service and the outside option is given by $\mathcal{U}^i(p, e_i, b_i) = \mathcal{U}^o(b_i)$ or else:

$$b_i(s,p) \equiv \frac{le_i\phi + (1-e_i)(p-\Delta y)}{(1-e_i)\Delta l + l - 1}, \tag{10}$$

with $l = 0$ when the consumer is compliant ($i = c$) and $l = 1$ if the consumer is not compliant ($i = nc$). The marginal compliant consumer is given by

$$b_c(s,p) = (1 - e_c(s))(\Delta y - p), \tag{11}$$

whereas the marginal non-compliant consumer is given by

$$b_{nc}(s,p) = \frac{p}{\Delta} - y + b_\phi(s), \tag{12}$$

with

$$b_\phi(s) \equiv \frac{\phi e_{nc}(s)}{\Delta(1 - e_{nc}(s))}. \tag{13}$$

If $e_i = 1$, the consumer obtains the utility of purchasing the lower-quality outside option and there is no demand for the higher-quality service from consumers in category $i \in \{c, nc\}$.

If the market is covered, the type of consumer indifferent between being compliant and not being compliant when she consumes the higher-quality service is given by $\mathcal{U}^c(p, e_c, b_I) = \mathcal{U}^{nc}(p, e_{nc}, b_I)$ or else:

$$b_I(s,p) \equiv \frac{(\Delta y - p)(e_{nc}(s) - e_c(s)) + \Delta(1 - e_{nc}(s))b_\phi(s)}{1 + \Delta(1 - e_{nc}(s))}. \tag{14}$$

The definitions of indifferent consumers given in (11), (12) and (14), respectively, enable us to determine in Lemma 1 the demand of each category of consumers for the higher-quality service.

Lemma 1. *If $p > \Delta y$, the monopoly excludes compliant consumers and we have $\mathcal{B}_c = \emptyset$ and $\mathcal{B}_{nc} = (min\{b_{nc}, B\}, B)$.*

If $\underline{p}(s) \le p \le \Delta y$, *with*

$$\underline{p}(s) \equiv \Delta y - \frac{\Delta b_\phi(s)}{1 + \Delta(1 - e_c(s))},$$

the market is not covered and we have $\mathcal{B}_c = (0, max\{b_c, 0\})$ *and* $\mathcal{B}_{nc} = (min\{b_{nc}, B\}, B)$.

If $p < \underline{p}(s)$, *the market is covered and we have* $\mathcal{B}_c = (0, max\{b_I, 0\})$ *and* $\mathcal{B}_{nc} = (min\{b_I, B\}, B)$.

For all $p \in \mathbb{R}_+$ *and* $e_i \in [0,1]$, *the demand* D_i *of each category* $i \in \{c, nc\}$ *of consumers decreases with the price* p *and the probability* e_i *of being excluded.*

Proof. See Appendix A-1. □

The choice of p and s determines whether the monopoly either serves both consumer categories or only one of them. The monopoly can exclude compliant consumers by choosing a very high price for its service, but this situation may not socially desirable if some non-compliant consumers generate social damage.

Lemma 2 shows that the complete exclusion of non-compliant consumers is possible if the price is higher than a minimal value that depends on the quality of the selection technology.

Lemma 2. *For a given quality* $s \in [\overline{s}, S)$ *of the selection technology, there exists a minimum price* $\hat{p}(s) \equiv \Delta\big(B + y - b_\phi(s)\big)$ *such that the monopoly completely excludes non-compliant consumers if it chooses* $p \ge \hat{p}(s)$. *The minimum price* $\hat{p}(s)$ *decreases with the quality* s *of the selection technology and the fine* ϕ *that is imposed on non-compliant consumers.*

Proof. See Appendix A-2. □

Sorting consumers by technology is required to exclude non-compliant consumers. However, even if their complete exclusion is technically possible, this strategy may be too costly for the monopoly, both in terms of selection costs and foregone earnings. Moreover, the exclusion of non-compliant consumers may have an ambiguous effect on the probability of excluding compliant ones. In practice, for example, the monopoly may sometimes make more type I errors and exclude a higher volume of compliant consumers if the quality of the selection technology is higher (but not maximal).[25] Therefore, in the following, we determine the monopoly's profit-maximizing selection strategy implied by the choice of s and p.

There are three possible strategies $j \in \{ne, dr, ib\}$, which imply that the monopoly may either serve one or two consumer categories:

- **a no-exclusion strategy ($j = ne$):** a high price, which implies that there is no demand from compliant consumers, and no adoption of the selection technology,
- **a de-risking strategy ($j = dr$):** a selection technology and a price, which imply that there is no demand from non-compliant consumers,
- **an imperfect blocking strategy ($j = ib$):** a selection technology and a price, which may generate a positive demand from both compliant and non-compliant consumers, respectively.

We denote the monopoly's maximum profit of choosing strategy $j \in \{ne, dr, ib\}$ by $\tilde{\pi}^j$, and it is obtained with a quality s^j for the selection technology and a price p^j for the service. The monopoly's maximum profit is given by $\tilde{\pi}^m = \max\{\tilde{\pi}^{ib}, \tilde{\pi}^{ne}, \tilde{\pi}^{dr}\}$, where $m \in \{ne, dr, ib\}$ represents the monopoly's profit-maximizing strategy.

We formalize an additional assumption, which implies (A3), to ensure that the monopoly has a positive demand from compliant consumers with the de-risking strategy:

(A6) $e_0 \leq \frac{\Delta B}{\phi + \Delta B}$.

Lemma 3 derives the conditions on the quality of the selection technology so that the monopoly adopts the de-risking strategy or the imperfect blocking strategy.

Lemma 3. *i) If $b_\phi(s) < B$, the de-risking strategy is not profitable. The monopoly trades off between choosing the imperfect blocking strategy and the no exclusion strategy.*

ii) If $b_\phi(s)(1 - \frac{1}{1+\Delta(1-e_c(s))}) \leq B \leq b_\phi(s)$, the monopoly trades off between choosing the de-risking strategy with $\hat{p}(s) \leq p \leq \Delta y$, the imperfect blocking strategy with $\underline{p}(s) \leq p \leq \hat{p}(s)$, and the no-exclusion strategy. In this case, with the imperfect blocking strategy, the market is not covered.

[25] Recall that we did not assume that the probability of excluding compliant consumers e_c decreases with the quality of the selection technology.

iii) If $B < b_\phi(s)(1 - \frac{1}{1+\Delta(1-e_c(s))})$*, the monopoly trades off between choosing the de-risking strategy with* $\hat{p}(s) \leq p \leq \underline{p}(s)$*, the imperfect blocking strategy with* $0 \leq p \leq \hat{p}(s)$*, and the no-exclusion strategy. In this case, with the imperfect blocking strategy, there is full market coverage.*

Proof. From Lemma 1 and Lemma 2, if $\hat{p}(s) > \Delta y$, both categories of consumers opt out the higher-quality market if the monopoly adopts the de-risking strategy. Therefore, the monopoly cannot make a positive profit with the de-risking strategy. Points ii) and iii) of Lemma 3 result from the comparison of the limit de-risking price $\hat{p}(s)$ with $\underline{p}(s)$ given in Lemma 1. □

In the rest of the paper, we choose to focus on exclusion probabilities e_c and e_{nc} and a parameter $B > 0$ so that the market is not covered with the imperfect blocking strategy, that is, case ii) of Lemma 3. In this case, the choice of the performance of the selection technology can imply the exclusion of some consumers from the higher-quality market.

In what follows, we denote by $\tilde{\pi}_s^j = \left|\partial\tilde{\pi}^j/\partial s\right|_{(p^j,s^j)}$ and $\tilde{\pi}_p^j = \left|\partial\tilde{\pi}^j/\partial p\right|_{(p^j,s^j)}$ the partial derivatives of the monopoly's profit with respect to s and p and $\tilde{\pi}_{sp}^j = \left|\partial^2\tilde{\pi}^j/\partial s\partial p\right|_{(p^j,s^j)}$ the cross derivative of the monopoly's profit with respect to s and p evaluated at the profit-maximizing price and quality, respectively.

4.2 The no-exclusion profit

If the monopoly chooses to only serve non-compliant consumers, its profit is decreasing with the quality of the selection technology because the demand of non-compliant consumers decreases with the probability of being excluded. Therefore, the monopoly prefers not to select its consumers at all and incurs the marginal cost $k_{nc}(0) = r_{nc}$. It chooses the price p^{ne} that maximizes its profit, sells its service to a quantity q^{ne} of non-compliant consumers and makes a profit $\pi^{ne} = \pi^{nc}(0, q^{ne})$.

4.3 The quality of the selection technology with imperfect blocking

Suppose that the monopoly prefers to adopt an "imperfect blocking" strategy with an uncovered market. From Lemma 3, this means that the profit-maximizing price and quality of the selection technology (p^{ib}, s^{ib}) are chosen so that $\underline{p}(s^{ib}) < p^{ib} < \hat{p}(s^{ib})$. At (p^{ib}, s^{ib}), the second-order conditions of profit-maximization must be locally satisfied, that is, it must be that $\tilde{\pi}_s^{ib} < 0$, $\tilde{\pi}_p^{ib} < 0$ and $\delta \equiv \tilde{\pi}_s^{ib}\tilde{\pi}_p^{ib} - \tilde{\pi}_{sp}^{ib} > 0$.[26]

Proposition 1 extends the result of Veiga and Weyl (2016) to our setting, in which the monopoly chooses the quality of a selection technology.

Proposition 1. *If the monopoly prefers the imperfect blocking strategy and the market is not covered, the profit-maximizing price and quality of the selection technology are implicitly defined by*

$$M\mathbb{E}\left[(1-e)\pi\,|\mathcal{M}\right] - \mathbb{E}\left[(1-e)\,|\mathcal{B}\right] = 0,$$

and

$$\begin{aligned}&-\mathbb{E}\left[(1-e)k' + em'\,\middle|\mathcal{B}\right] + \mathbb{E}\left[\tilde{u}' + e'p\,\middle|\mathcal{M}\right]\mathbb{E}\left[1-e\,|\mathcal{B}\right]\\ &+\frac{\text{Cov}\left[\tilde{u}' + e'p, \pi\,|\mathcal{M}\right] + \mathbb{E}(e\pi\,|\mathcal{M})\,\mathbb{E}(\tilde{u} + e'p\,|\mathcal{M})}{MS} - \mathbb{E}\left[e'(P-k+m)\,\middle|\mathcal{B}\right] = 0,\end{aligned}$$

with $\pi = (1-e)(P-k) - em$.

Proof. See Appendix B-1. □

An increase in the quality of the selection technology has four effects on the firm's profit.

1. **A direct cost effect**: when the firm increases the quality of the selection technology, it loses the average increase in the cost of serving consumers, which results from an increase in the quality of the selection process, $\mathbb{E}\left[(1-e)k' + em'\,|\mathcal{B}\right]$. In particular, in our setting, since $k' = m' = a'$, we have $\mathbb{E}\left[(1-e)k' + em'\,|\mathcal{B}\right] = \mathbb{E}\left[a'\,|\mathcal{B}\right]$.

2. **A private sorting effect**: a higher quality of the selection technology has a sorting effect on the firm's profit, which depends on whether the marginal consumers who are most strongly attracted by a better-quality for the selection technology are those who are more costly to

[26]If there is a corner solution such that one category of consumers is excluded, the monopoly trades off between the strategies "no-exclusion" and "de-risking". We discuss in the Appendix the conditions under which there may be an interior solution.

serve.[27]

3. **A private exclusion effect**: when the firm increases the quality of the selection technology, this impacts the marginal probability that a buyer is excluded from the market. If the probability of a consumer being excluded increases on average, this represents a cost to the firm, $\mathbb{E}\left[e'(P-k+m)\,|\mathcal{B}\right]$, that is, the cost of foregone earnings $P-k$ and the additional cost of missed sales m.

4. **A private Spence term**: the firm raises the price by $\mathbb{E}\left[\tilde{u}'+e'p\,|\mathcal{M}\right]\mathbb{E}\left[1-e\,|\mathcal{B}\right]$ when it increases the quality s of the selection technology, because to hold fixed q, the price must offset the average benefit that marginal consumers derive from the variation of the quality of the selection technology, which affects the probability that they are excluded.

Veiga and Weyl (2016) do not consider the probability that a consumer may be excluded and obtain the same effects (1),(2), and (4) as $e=0$. The private exclusion effect is specific to our model.

A special case with $a_i(s)=0$ and $t_c=t_{nc}=0$:
Proposition 2 gives the choice of the price of the service and the quality of the selection technology in the special case where there is no cost to collect consumer data (i.e, $a_i(s)=0$) and there are no additional costs to exclude consumers (i.e, $t_c=0$). In this case, the marginal cost of serving a consumer in category $i\in\{c,nc\}$ is $k_i(s)=r_i$ for $i\in\{c,nc\}$, with $r_c=0$.

To proceed, it is convenient to denote the expected demand of consumers in category $i\in\{c,nc\}$ who are not excluded by $\tilde{D}_i\equiv(1-e_i)D_i$. We also denote by $\tilde{\eta}_i=-(\partial\tilde{D}_i/\partial s)(s/\tilde{D}_i)$ the elasticity of $\tilde{D}_i$ to the choice of the quality of the selection technology.

Proposition 2. *If the monopoly prefers the imperfect blocking strategy, if there are no costs of data collection and no additional costs incurred for the exclusion of consumers, the profit-maximizing price and quality of the selection technology are chosen such that*

$$\frac{\tilde{D}_c p^{ib}}{\tilde{D}_{nc}(p^{ib}-r_{nc})}=-\frac{\tilde{\eta}_{nc}}{\tilde{\eta}_c},\tag{15}$$

[27] We refer the reader to Veiga and Weyl (2016) for the analysis of the distinction between selection and sorting in the literature (Akerlof, 1970, Einav and Finkelstein, 2011).

and for $i \in \{c, nc\}$

$$\sum_{i \in \{c,nc\}} \big((1 - e_i(s^{ib}))M_i(p^{ib} - k_i(s^{ib})) - \tilde{D}_i\big) = 0. \tag{16}$$

Proof. See Appendix B-2. □

If the monopoly prefers the imperfect blocking strategy, the ratio of its profit from each consumer category depends on the relative elasticities of consumer demands that are not excluded in each category, respectively.[28] This result resembles the pricing formula obtained in the literature on platform markets for monopolistic platforms (e.g. Rochet and Tirole (2003)). In particular, Eq.(15) represents the relative mark-up in each category of consumers, which is similar to the role of the price structure in two-sided markets, when, ex ante, price discrimination is possible. In our setting, the choice of the quality of the selection technology creates an externality between compliant and non-compliant consumers, respectively. In contrast, in the two-sided markets literature, there is a direct network externality between different categories of users.[29]

Sanctioning mis-classification type I and type II errors:

Regulators often focus on increasing sanctions for type II errors, when the monopoly erroneously sells to non-compliant consumers.[30] Actually, such sanctions may not always increase the monopoly's incentives to select its consumers by technology. To see why, we denote by $p^{ib}_{r_{nc}}$ and $s^{ib}_{r_{nc}}$ the marginal impact of an increase in r_{nc} the sanction for type II errors on the price of the service and the quality of the selection technology, respectively. If $a_c = a_{nc} = 0$, applying the implicit function theorem to

[28] If $\tilde{\eta}_{nc} \geq 0$, which happens when e_{nc} is increasing, a necessary condition for an interior solution to exist is that there exist a price p and a quality s for the selection technology such that $\tilde{\eta}_c \leq 0$, because we consider that the social planner cannot subsidize sales with negative sanctions.

[29] Therefore, our setting differs from Weyl (2010) of monopolistic platform pricing, because we relax assumption (ii) of his paper, that is, all groups can be price discriminated, and assumption (iv), that is, externalities are only to participating users. Weyl (2010) clearly mentions that assumption (iv) is not necessary to his result and that it is connected to the literature on contracting with externalities (Segal, 1999).

[30] In our setting, the regulator is systematically able to detect the monopoly's mis-classification errors and sanction them. In the case of Anti-Money Laundering, see the contribution by Takáts (2011). Banks face fines if they fail to report suspicious transactions. In this setting, where the precision of the signal is given, banks tend to report transactions excessively in order to avoid being punished as fines increase. Reporting depends on both monitoring effort and the reporting threshold. As fines increase, the reporting threshold decreases, thereby deteriorating the quality of the information in the report. The author's stylized facts support the theoretical predictions. In contrast, in our setting, fines may affect the quality of the selection technology.

the first-order conditions of profit-maximization yields:

$$p^{ib}_{r_{nc}} = \frac{1}{\Delta\delta}(-(1-e_{nc})f(b_{nc})\tilde{\pi}^{ib}_{s} + \Delta D^{'}_{nc}e^{'}_{nc}\tilde{\pi}^{ib}_{sp}), \tag{17}$$

and

$$s^{ib}_{r_{nc}} = \frac{1}{\Delta\delta}(-\Delta D^{'}_{nc}e^{'}_{nc}\tilde{\pi}^{ib}_{p} + (1-e_{nc})f(b_{nc})\tilde{\pi}^{ib}_{sp}), \tag{18}$$

where $D^{'}_{nc}e^{'}_{nc} \leq 0$, $\delta > 0$, $\tilde{\pi}^{ib}_{p} < 0$ and $\tilde{\pi}^{ib}_{s} < 0$. The signs of $p^{ib}_{r_{nc}}$ and $s^{ib}_{r_{nc}}$ depend on the sign of the cross-derivative of the profit with respect to s and p and the optimum, that is, $\tilde{\pi}^{ib}_{sp}$. In Appendix B-4, we show through numerical simulations that the sign of $\tilde{\pi}^{ib}_{sp}$ varies according to the choice of exclusion probabilities.

If $\tilde{\pi}^{ib}_{sp} < 0$, a higher quality of the selection technology reduces the marginal profit that the monopoly obtains when it increases its price. In this case, the price of the service and the quality of the selection technology are substitute strategies to maximize the monopoly's profit. Then, from (17) and (18), a higher sanction for type II errors increases the monopoly's price and lowers the quality of the selection technology; that is, we have $p^{ib}_{r_{nc}} \geq 0$ and $s^{ib}_{r_{nc}} \leq 0$.

If $\tilde{\pi}^{ib}_{sp} > 0$, a higher quality of the selection technology gives the monopoly incentives to increase its price. Therefore, the price of the service and the quality of the selection technology are complements. Then, a higher sanction for type II errors has an ambiguous impact on the price of the service and the quality of the selection technology. For example, if e_{nc} is a constant and $e^{'}_{nc} = 0$, in this case, a higher sanction for type II errors provides the monopoly with incentives to increase its price and the quality of the selection technology; that is, we have $p^{ib}_{r_{nc}} \geq 0$ and $s^{ib}_{r_{nc}} \geq 0$.

In Appendix B-4, we illustrate cases in which the quality of the selection technology is increasing or decreasing with the sanction for type II errors. We also perform the same analysis with respect to the sanction for type I errors t_c and the fine ϕ imposed on non-compliant consumers.

We conclude that higher sanctions for type I and type II errors do not systematically result in an increase in the quality of the selection technology, nor do they necessarily imply higher prices.[31]

[31] This result is reminiscent of the paper by Jarle Kind et al. (2008), which shows that taxation in two-sided markets may sometimes result in an increase in the production of the goods offered to each side of the market.

Depending on consumer demand response to variations in the quality of the selection technology and the sign of $\tilde{\pi}^{ib}_{sp}$, both instruments have sometimes opposite effects on the monopoly's choice of the quality of a selection technology, whereas they may sometimes reinforce each other.

4.4 The de-risking profit

The monopoly may decide to try and exclude all non-compliant consumers from the market. To that end, according to Lemma 3, it should choose $p \in [\hat{p}(s), \underline{p}(s)]$.[32] If there is an interior solution to the monopoly's profit-maximization problem, there are the same effects as in section 4.3 when two consumer categories are served.[33] The monopoly makes a profit $\tilde{\pi}^{dr} = \pi^c(s^{dr}, p^{dr})$. If the constraint $p \geq \hat{p}(s)$ is binding, the monopoly chooses the price $\hat{p}(s)$ given in Lemma 2 and the quality s of the selection technology maximizes $\pi^c(s, \hat{p}(s))$. We assume that the constrained profit of the monopoly is concave with respect to s.

Proposition 3 gives the profit-maximizing price when the monopoly is constrained in its pricing strategy by the choice to exclude non-compliant consumers.

Proposition 3. *If the monopoly wishes to completely exclude non-compliant consumers and is constrained when it maximizes its profit, it chooses a price and a quality of the selection technology such that $\hat{p}(s^{dr}) = \Delta\big(B + y - b_\phi(s^{dr})\big)$, and*

$$\frac{\hat{p}(s^{dr}) - a_c(s^{dr})}{\hat{p}(s^{dr})} = \frac{1}{\epsilon_c} - \mu^{dr}, \tag{19}$$

with ϵ_c the price elasticity of the demand of compliant consumers, and

$$\mu^{dr} = \frac{(\pi^c)'|_{s=s^{dr}}}{\hat{p}(s^{dr})B\Delta\left(D'_{nc}\right)|_{s=s^{dr}}}.$$

Provided that the monopoly makes a positive margin, it makes a profit $\pi^{dr} = \pi^c(s^{dr}, \hat{p}(s^{dr}))$, and makes no profit otherwise.

[32] Recall that we focus on case ii) of Lemma 3.

[33] The only difference comes from the fact that there is no demand from non-compliant consumers, which implies that the set of buyers is $\mathcal{B}_c$ instead of $\mathcal{B}$, the set of marginal buyers is $\mathcal{M}_c$ instead of $\mathcal{M}$, and that the sorting effect is calculated when there is no demand from non-compliant consumers.

Proof. See Appendix B-3. □

If the monopoly is constrained by its choice to exclude non-compliant consumers, it compensates compliant consumers by choosing a lower price or extracts more surplus from them. Compared to the standard Lerner pricing formula, there is an additional price reduction (if μ^{dr} is positive) or mark-up (if μ^{dr} is negative) that is caused by the choice to exclude non-compliant consumers. To understand why, suppose that the cost of collecting consumer data is null (i.e. $a_c(s) = 0$), there is no cost of excluding a compliant consumer (i.e. $t_c = 0$) and e_{nc} is increasing with s. Since $(\pi^c)'|_{s=s^{dr}} = \hat{p}(s^{dr})\ \tilde{D}_c'\Big|_{s=s^{dr}}$, we have the following:

$$\mu^{dr} = \frac{\tilde{D}_c'\Big|_{s=s^{dr}}}{B\Delta\left(D_{nc}'\right)|_{s=s^{dr}}}.$$

Since $D_{nc}' < 0$ from Lemma 1 and $B\Delta > 0$, the term μ^{dr} is positive if, at the profit-maximizing price and quality, the expected probability of selling to compliant consumers $\tilde{D}_c$ decreases with the quality of the selection technology. To exclude non-compliant consumers, the monopoly needs to increase the quality of the selection technology. If this has a negative effect on the expected probability of selling to compliant consumers, the monopoly compensates the latter for the negative externality exerted by the non-compliant consumers, who are excluded. Therefore, it offers them a lower price.[34] The price decrease is stronger when the demand of non-compliant consumers is not very responsive to the choice of the quality of the selection technology, while the expected probability of selling to a compliant consumer is very responsive to the quality of the selection technology.

4.5 The monopoly's optimal strategy

The monopoly compares its maximum profit of serving either one category of consumers or both. Without any sanctions, the monopoly sometimes makes a higher profit when it does not exclude

[34]This result is similar to other environments in which there is contracting with externalities between participating and non-participating consumers (see Segal, 1999). In a related literature, Damiano and Li (2007) study competition between matchmakers, when agents have heterogeneous qualities and the expected quality of the pool of participants affects a user's decision to join an intermediary.

any consumer than with the de-risking strategy. However, this is not systematic because of different effects. First, excluding consumers is costly, which makes the “no-exclusion” strategy more profitable than “de-risking”. Second, non-compliant consumers have a higher willingness to pay for the monopoly’s service, and therefore, it is more profitable to serve them. However, the price elasticity of both consumer categories differs. With the de-risking quality for the selection technology, from (10), the marginal non-compliant consumer is more sensitive to prices than the marginal compliant one if and only if $1/\Delta \geq (1 - e_c(s^{dr}))$. When the marginal non-compliant consumer is relatively more sensitive to prices, there is a countervailing effect, which increases the monopoly’s incentives to serve compliant consumers rather than non-compliant ones.

Figures 1-a and 1-b below illustrate the monopoly’s optimal strategy for the parameter values given in Table 0 of Appendix A-5 in two different cases, when the quality of the selection technology is respectively decreasing or increasing with the sanction for type II errors. We see for example in Figure 1-a that if the regulator chooses a high sanction for type II errors, the monopoly only sells to non-compliant consumers, because the sanction is passed through to consumers into higher prices, which implies the exclusion of compliant consumers. In figure 1-b, we see that the monopoly sells to both types of consumer for low values of the sanction, then chooses the de-risking strategy when the sanction increases, and only sells to non-compliant consumers for high values of the sanction.

Figure 1-a: The monopoly’s strategy when s^{ib} is decreasing with the sanction r_{nc}

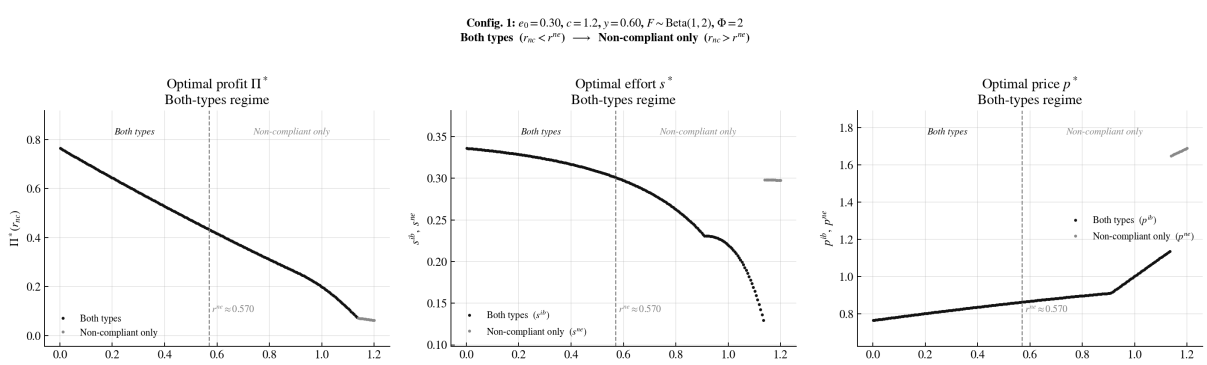
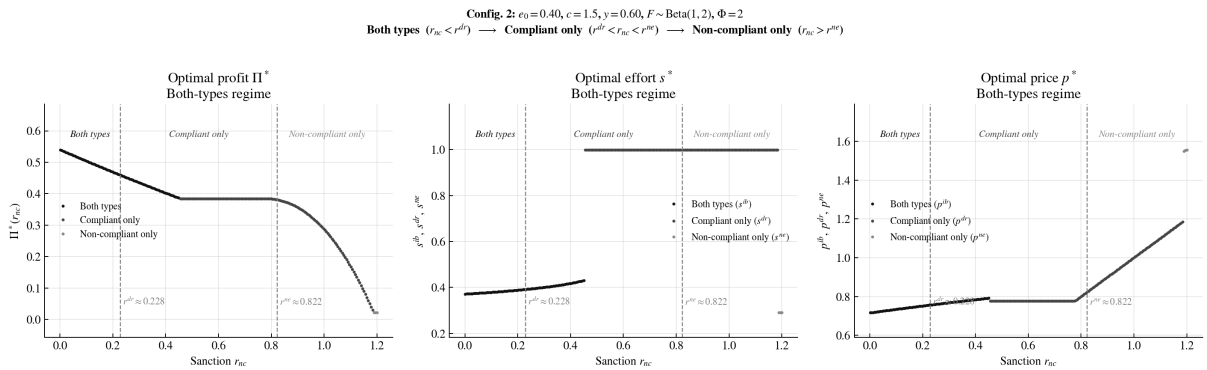


Figure 1-b: The monopoly's strategy when s^{ib} is increasing with the sanction r_{nc}



5 Regulation of the selection technology

In this section, we start by determining the first-best situation, when the social planner is able to control both the choice of the quality of the selection technology and the price of the service.[35] We compare the firm's private choice of the quality of its selection technology to the first-best.

Then, we analyze a second-best situation, when the social planner can indirectly impact the firm's selection costs through sanctions for mis-classification errors.[36] We also compare several instruments that the regulator can use to improve the allocation: a price cap $\overline{p}$ and a variation in the fine ϕ imposed on non-compliant consumers.

5.1 Welfare maximization

We start by determining the first-best choice of the quality of the selection technology and the price and explain why the monopoly's choice is not socially optimal. We formalize another assumption:

(A5) if the monopoly adopts the no-exclusion strategy with $s = 0$, $e_0 = 0$ and $p = p^{ne}$, social welfare is reduced relative to a situation in which the monopoly does not sell the higher-quality service, that is, we have $W(0, p^{ne}) < \underline{W}$.

[35] The regulator may also be able to control ϕ the fine imposed on non-compliant consumers. This interpretation is plausible in contexts where the regulatory authority has control over audit intensity or monitoring resources, which is common in prudential, environmental, and tax enforcement settings. However, this assumption is less plausible if this parameter reflects technological or judicial constraints that are exogenous to the regulator.

[36] Naturally, the possibility for the regulator to use each type of tool depends on the observability of the monopoly's choices. A monopoly's incentives to misreport its selection technology is analyzed by Colliard (2019).

Therefore, the social planner trades off between a de-risking strategy, or an imperfect blocking strategy. The maximum welfare is $W^w = \max(W^{ib}, W^{dr})$, where $w \in \{dr, ib\}$ represents the social planner's welfare-maximizing strategy.[37]

If the monopoly chooses to sell to the socially optimal set of consumer categories, it chooses neither the welfare-maximizing quality of the selection technology nor the welfare-maximizing price. The sorting and the Spence distortions are standard when a monopoly is able to jointly choose a price and the quality of a service (see Spence (1975) and Veiga and Weyl (2016)). In our setting, there is also an *exclusion distortion* because the monopoly's strategy generates an inefficient relative participation of each consumer category in the higher-quality market compared to the first-best. This distortion arises even when both the monopoly and the social planner prefer to serve both categories of consumers. Sometimes, the exclusion distortion implies that the monopoly does not sell to the socially optimal set of consumer categories. For example, it may choose to sell to both categories of consumers or adopt a no-exclusion strategy, though a social planner prefers the de-risking strategy with a complete exclusion of non-compliant consumers. Or, it may choose to adopt a no-exclusion strategy, while the social planner prefers either a de-risking strategy or an imperfect blocking strategy.

5.1.1 Social welfare with an imperfect blocking strategy

Suppose that both the monopoly and the social planner prefer the monopoly to serve both categories of consumers. Proposition 4 gives the first-best choice of the price of the service and the quality of the selection technology.

Proposition 4. *Suppose that the social planner prefers the imperfect blocking strategy. The welfare-*

[37] There are contexts in which having no model at all can be better than having a model. For example, Sarlin (2013) introduces a measure of the model's usefulness for early warning systems of crisis detection, comparing the loss incurred by a social planner when it does not use a model and when it uses one. The loss function of his paper accounts for class specific and observation specific misclassification errors.

maximizing quality of the selection technology and price are implicitly defined by

$$
\begin{aligned}
&- \mathbb{E}\left[(1-e)k' + em' \,\middle|\, \mathcal{B}\right] + \frac{\operatorname{Cov}\left[\tilde{u}' + e'p, \tilde{w} \,|\, \mathcal{M}\right] + \mathbb{E}(e\tilde{w} \,|\, \mathcal{M})\, \mathbb{E}(\tilde{u} + e'p \,|\, \mathcal{M})}{MS} \\
&- \mathbb{E}\left[e'(w - k - o + m) \,\middle|\, \mathcal{B}\right] = 0,
\end{aligned}
$$

and

$$
\mathbb{E}\left[(1-e)\tilde{w} \,|\, \mathcal{M}\right] = 0,
$$

with $\tilde{w} = (1-e)(w-k-o) - em$.

Proof. See Appendix C-1. □

An increase in the quality of the selection technology has three effects on social welfare.

1. **A direct cost effect**: when the firm increases the quality of the selection technology, it loses the average increase in the cost of all consumers, which results from an increase in the quality of the selection technology, $\mathbb{E}\left[(1-e)k' + em' \,|\, \mathcal{B}\right]$. This term is identical to the private direct cost effect for a given price. In particular, since $k' = m' = a'$ in our setting, the firm loses $\mathbb{E}\left[(1-e)k' + em' \,|\, \mathcal{B}\right] = \mathbb{E}\left[a' \,|\, \mathcal{B}\right]$.

2. **A social sorting effect**: a higher quality of the selection technology has a sorting effect on social welfare, which depends on whether the marginal consumers who are most strongly attracted by a better-quality for the selection technology are those who generate a higher social benefit of being served by the firm. In our setting, the contribution of marginal consumers to welfare differs from their contribution to profit, even if the monopoly chooses the imperfect blocking strategy. The difference between the private sorting effect and the social sorting effect generates the sorting distortion.

3. **A social exclusion effect**: when the social planner increases the quality of the selection technology, this impacts the marginal probability that a buyer is excluded from the market. If the probability that a consumer is excluded increases in average, this represents a change in the average social welfare, $\mathbb{E}\left[-e'(w-k-o+m) \,|\, \mathcal{B}\right]$. The difference between the private

exclusion effect and the social exclusion effect generates the *exclusion distortion*. In our setting, we have $\mathbb{E}\left[-e'(w-k-o+m)\,|\mathcal{B}\right] = \mathbb{E}\left[-e'(w-o+\lambda(t-r))\,|\mathcal{B}\right]$.

In Veiga and Weyl (2016), there are the effects (1) and (2) with $e = 0$, but not the effect (3). Moreover, in their model, there is a social Spence term, which is absent in our setting because the quality of the service is exogenous. The social planner internalizes the preferences of all buyers, while the monopolist only internalizes the preferences of the marginal buyers when it raises its price.[38]

5.1.2 Social welfare with a de-risking strategy

When the fine imposed on non-compliant consumers is high or when non-compliant consumers cause a high social damage, the de-risking strategy is socially optimal. In this case, from (4), social welfare can be expressed as

$$W(s,p) = \mathbb{E}\left[(1-e)\Delta y - \lambda e t - a\,|\mathcal{B}_c\right] + \mathbb{E}(o(\theta)).^{39} \tag{20}$$

Because compliant consumers do not generate social damage, the de-risking welfare increases with their participation in the higher-quality market. Therefore, the social planner chooses the minimum price so that non-compliant consumers are excluded, that is, $\hat{p}(s)$, which is given in Lemma 2. Its choice of the quality of the selection technology s^w differs from that of the monopoly due to the effects identified in Proposition 4. In this case, the direct cost effect is equal to $-\mathbb{E}\left[a'\,|\mathcal{B}_c\right] = -a_c'(s^w)F(b_c) < 0$ because there is only a demand from compliant consumers. The social exclusion effect is equal to $-\mathbb{E}\left[e'(w-o+\lambda(t-r))\,|\mathcal{B}_c\right] = -e_c^{'}(s^w)(\Delta y+\lambda t_c)F(b_c)$ and it is negative if $e_c^{'}(s) > 0$. The social sorting effect depends on the price that is chosen such that non-compliant consumers are excluded and it is equal to

$$(D_c^{'} - B\Delta(1-e_c(s^w))D_{nc}^{'}f(b_c))\tilde{w}_c(s^w,p^w). \tag{21}$$

[38] In our setting, we have that $\mathbb{E}\left[w'\,|\mathcal{B}\right]$=0, which suppresses the role of the Social Spence term to simplify.

[39] We suppress r which is equal to zero for compliant consumers and use the assumption that compliant consumers do not generate social damage.

The sign of the social sorting effect given in Eq.(21) depends on the relative elasticities of the demands of non-compliant consumers and compliant consumers to the quality of the selection technology, respectively.[40] For example, if the demand of non-compliant consumers is very responsive to the choice of the quality of the selection technology, while it does not impact much the probability to exclude compliant ones, the sign of the social sorting effect is positive. In contrast, if the demand of non-compliant consumers is not very sensitive to the choice of the quality of the selection technology, while a higher quality of the selection technology increases the probability of excluding compliant consumers, the sign of the social sorting effect may sometimes become negative. There are two cases:[41]

- If the social sorting effect is negative (or positive but of low magnitude), social welfare can be decreasing with the quality of the selection technology if $e_c^{'} > 0$ or if e_c is not very responsive to variations in s. In this case, the social planner chooses the minimal quality of the selection technology that is compatible with the exclusion of non-compliant consumers and the participation of the monopoly.

- If the social sorting effect is positive, the social planner chooses a quality for the selection technology that reflects a trade-off between the social cost of exclusion, the direct cost of selecting consumers, and the marginal social benefits of having more compliant consumers who are included in the higher-quality market.

5.1.3 Comparing the welfare effects of selection technologies

The social planner chooses the welfare-maximizing strategy w, which implies that there is a set $\mathcal{B}_i^w(s^w, p^w)$ of buyers in category $i \in \{c, nc\}$. The monopoly chooses the profit-maximizing strategy m, which implies that there is a set $\mathcal{B}_i^m(s^m, p^m)$ of buyers in category $i \in \{c, nc\}$.

Proposition 5 evaluates the effect of the monopoly's choice of a quality for the selection technology and a price for the service with respect to the first-best.

[40] Recall that we have $D_c^{'} = -e_c^{'} f(b_c)(\Delta y - p)$ and $D_{nc}^{'} < 0$ if $e_{nc}^{\prime} > 0$.

[41] Recall that we have $b_\phi(s^w) - B \geq 0$ and $D_{nc}^{'} < 0$.

Proposition 5. *With respect to the first-best, the adoption of at selection technology of quality* s^m *and with a price* p^m *for the service generates a welfare loss given by:*

$$L(s^m, p^m) = \mathbb{E}\left[\tilde{w}\,|\mathcal{B}^w(s^w, p^w)\right] - \mathbb{E}\left[\tilde{w}\,|\mathcal{B}^m(s^m, p^m)\right],$$

with $\tilde{w} = (1-e)(w - \lambda r - o) - \lambda e t - a$.

Proof. From (4), we have $W^w(s^w, p^w) - W^m(s^m, p^m) = L(s^m, p^m)$. □

The loss function measures how the imperfection of the selection technology impacts social welfare when there is both selection by quantity and technology. The loss function of Proposition 5 has some similarities with the literature on algorithms, which analyzes the social planner's regulation of selection decisions. However, this literature considers the choice of the performance of an algorithm for a given price. Therefore, it neglects how the *joint choice* of a performance for the algorithm and a price for a service impacts the endogenous proportion of buyers who belong to each category.[42]

The loss function can be decomposed as the sum of two terms. First, at the welfare-maximizing price p^w, there are the losses caused by the private choice of the quality of the selection technology, which is s^m instead of s^w. The usual evaluations of algorithmic performance focus on these losses, but do not mention which price is chosen as a reference point. Moreover, they often consider that the first-best is reached when there are no errors, whereas this may not be accurate when taking into account the costs of the selection process and the sorting effect. Second, at the profit-maximizing quality of the selection technology s^m, there are losses that are caused by the choice of the monopolistic price p^m instead of the welfare-maximizing price p^w. This term is generally overlooked in the usual evaluations of algorithmic performance.[43]

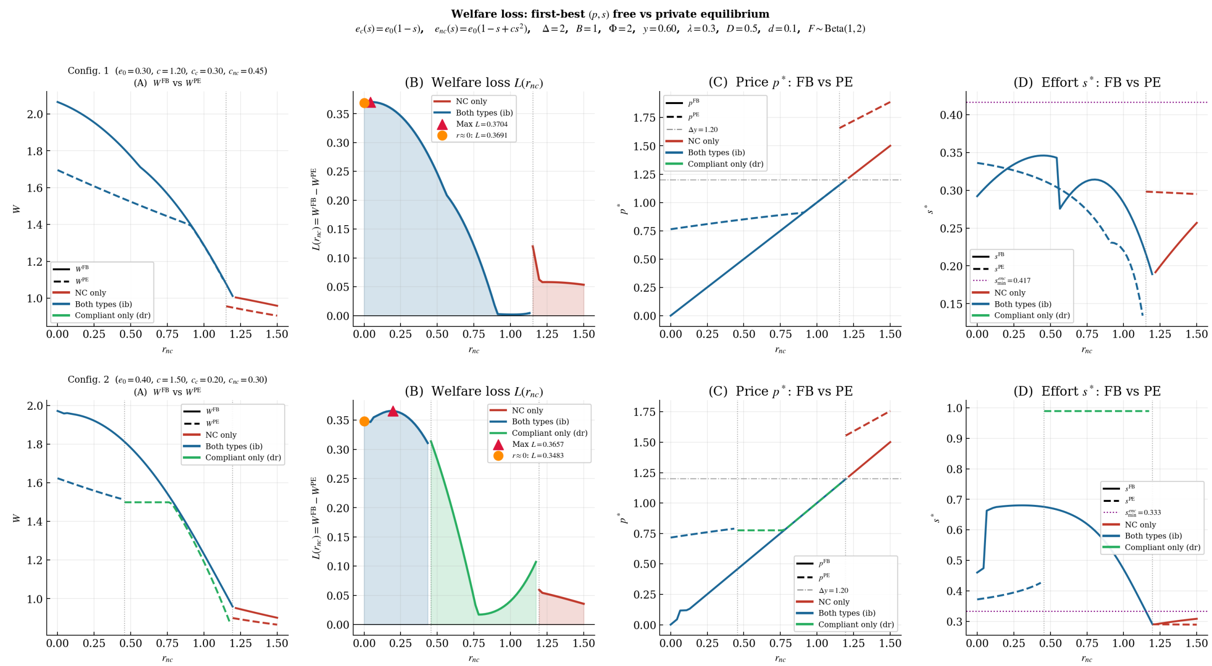
Figure 2 illustrates the welfare loss caused by the monopoly's choices according to the sanction for

[42] Rambachan et al. (2020) model the social preferences over the screening decisions as the sum of social weights multiplied by the expected average outcome of interest given a selection rule. They focus however on a different issue, that is, how a social planner should select a decision rule when for exogenous reasons, one group of users may be disadvantaged by an observable characteristic. In Section 4 of their paper, the social planner delegates the screening decision to a human decision-maker, whereas, in our paper, the decision-maker is a firm, and therefore, chooses a price for its service.

[43] See Appendix D.

type II errors with the parameters of Table 0 (see Appendix A-5).[44]

Figure 2: The welfare loss according to the sanction for type II errors



In the first configuration, the minimum welfare loss is $L^{\min} = 0.00197$ at $r_{nc} \approx 1.02$, and at that point both the regulator and the monopoly choose the same imperfect blocking strategy. Regime divergence only arises at the very end of the sanction range ($r \in [1.15, 1.22]$), where the monopoly switches to a no-exclusion strategy while the regulator maintains imperfect blocking. The minimum of the welfare loss corresponds to the point where the sanction r_{nc} constrains the monopoly's margin so tightly that its choices of (p^m, s^m) converge towards those of the regulator. In the second configuration, the minimum is $L^{\min} = 0.01718$ at $r_{nc} \approx 0.78$, and at that point the monopoly chooses the de-risking strategy while the regulator prefers imperfect blocking. It is therefore precisely in the zone of regime divergence that the welfare loss reaches its floor. As r_{nc} increases from 0.46 to 0.78, the monopoly abandons serving non-compliant consumers, setting a price $p^m \approx \Delta y/2$ that remains stable. In contrast, the regulator maintains imperfect blocking by lowering its price in order to retain non-compliant consumers. The two welfare curves converge

[44] In figure 2, FB stands for first-best, whereas PE stands for private equilibrium.

towards the minimum, after which W^m remains approximately constant in r_{nc}.[45]

5.2 Regulating with sanctions for mis-classification errors

In this section, we assume that the social planner can affect the cost of selecting different types of consumers through the choice of a sanction for each type of error, that is, t_c for type I errors, and r_{nc} for type II errors, respectively. Then, the monopoly responds by choosing the profit-maximizing strategy $m \in \{ne, ib, dr\}$, with the price $p^m(r_{nc}, t_c)$ and the quality of the selection technology $s^m(r_{nc}, t_c)$.

From assumption (A6), the second-best welfare-maximizing strategy w is either de-risking or imperfect blocking. To achieve it, the social planner must take into account that the choice of the sanctions for errors affects the monopoly's incentives to adopt the welfare-maximizing strategy. For each $w \in \{dr, ib\}$, the social planner maximizes

$$\max_{(r_{nc}, t_c)} \quad W$$

$$\text{s.t.} \quad \pi^w(p^w(r_{nc}, t_c), s^w(r_{nc}, t_c)) > \pi^j(p^j(r_{nc}, t_c), s^j(r_{nc}, t_c)) \quad \text{for } j \in \{ne, ib, dr\}, \text{ and } j \neq w \tag{IC}$$

$$\pi^w(p^w(r_{nc}, t_c), s^w(r_{nc}, t_c)) > 0. \tag{PC}$$

Then, the social planner compares the maximum achievable social welfare with the de-risking and the imperfect blocking strategy, given the constraints that the monopoly adopts the same strategy (IC) and participates in the market (PC).

5.2.1 Implementing the De-risking strategy

Implementing the de-risking strategy is always possible by imposing on the firm a sufficiently high sanction for each type II error. Such a sanction raises the monopoly's marginal cost of selling to non-compliant consumers with the imperfect blocking or the no exclusion strategy, respectively,

[45] With the de-risking strategy, the monopoly abandons serving non-compliant consumers, so neither its profit nor the social welfare evaluated at the monopoly equilibrium depends on r_{nc}. Any residual decline visible in the simulation reflects numerical approximation near the ib/dr regime boundary.

without affecting the de-risking profit.[46] The social planner chooses the sanction for type I errors that maximizes the de-risking welfare under the constrain that the monopoly participates in the market. Taking the total derivative of social welfare with respect to t_c yields

$$(W^{dr})'(t_c) = -e_c \lambda F(b_c) + s^{dr}_{t_c} \left(\frac{\partial W^{dr}}{\partial p} \frac{\partial \hat{p}}{\partial s} + \frac{\partial W^{dr}}{\partial s} \right),$$

where $s^{dr}_{t_c} > 0$ if $e_c D_c$ is decreasing with s. If the monopoly overinvests in the quality of the selection technology, the welfare-maximizing sanction for type I errors is equal to zero.[47] If the monopoly underinvests in the quality of the selection technology, having a strictly positive sanction for type I errors improves the allocation if social welfare is increasing at the de-risking profit-maximizing quality, which happens when the cost for the regulator of detecting mis-classification errors is sufficiently low.

5.2.2 Implementing the imperfect blocking strategy

If the social planner wishes to implement the imperfect blocking strategy, it needs to ensure that the monopoly actually prefers this strategy to no-exclusion or de-risking. However, given the presence of externalities between consumer categories, a sanction for type II errors may not be a sufficient instrument to achieve this objective. On the one hand, it is necessary to increase the sanction for type II errors to reduce the no-exclusion profit. On the other hand, with a higher sanction for type II errors, the de-risking strategy may sometimes become more attractive for the monopoly than imperfect blocking. Therefore, the social planner may have to use a second instrument to implement the imperfect blocking strategy.[48]

Such an implementation is sometimes possible, for example, by adding a sanction for type I errors

[46] The possibility to implement the de-risking strategy with sanctions for type II err rests on the assumption that the monopoly does not incur fixed entry costs. Would the monopoly incur fixed entry costs, the regulator would need to ensure its participation with fixed transfers, which would be potentially costly for the society.

[47] Recall that we assumed that W is concave with respect to s. If the monopoly over-invests in s, this implies that at $s = s^{dr}$, the derivative of social welfare with respect to s is negative. Thus, social welfare is decreasing with t_c.

[48] In the special case of anti-money laundering policies, this implementation problem could explain why public authorities have expressed concern that banks could inefficiently adopt de-risking strategies given the rise of the sanctions for type II errors.

(t_c in our model).[49] However, this may not be the best regulatory instrument for improving social welfare. The regulator could also decide to use a price cap $\overline{p}$ to prevent the monopoly from increasing its price, which affects its incentives to choose the de-risking strategy. Or it could reduce the level of the fine ϕ imposed on non-compliant consumers to ensure their participation in the market. However, both instruments have different effects on social welfare and the surplus of compliant consumers.

To understand how the choice of a regulatory instrument affects social welfare, we use the same notation σ for each of them. We also denote by p_σ^m and s_σ^m the marginal impact of an increase in σ on the price of the service and the quality of the selection technology, respectively, and by $k_\sigma = \partial k/\partial\sigma$ and $m_\sigma = \partial m/\partial\sigma$, respectively. Proposition 6 gives the per-consumer marginal effect of an increase in the regulatory instrument σ on social welfare:

Proposition 6. *The per-consumer marginal effect of an increase in the regulatory instrument σ on welfare is:*

$$\frac{\partial W}{\partial \sigma}\frac{1}{q} = -\mathbb{E}\left[(1-e)k_\sigma + em_\sigma \,|\mathcal{B}\right] - \mathbb{E}\left[(e'(w-\lambda(r-t)-o)+a')s_\sigma^m \,\middle|\mathcal{B}\right]$$

$$+\frac{1}{MS}(\mathbb{E}\left[(\tilde{u}'+e'p)\tilde{w}s_\sigma^m \,\middle|\mathcal{M}\right] + \mathbb{E}\left[(1-e)\tilde{w}p_\sigma^m \,|\mathcal{M}\right]),$$

with $k_{t_c} = 0$, $k_{\overline{p}} = 0$, $k_\phi = 0$,$k_{r_{nc}} = \lambda$, $m_{t_c} = \lambda$, $m_{\overline{p}} = 0$, $m_\phi = 0$ and $m_{r_{nc}} = 0$.

Proof. See Appendix E. □

The regulatory sanctions for errors increase the monopoly's selection costs, which may affect the prices paid by consumers and the quality of the selection technology. As shown in section 4.3, prices and the quality of the selection technology do not necessarily increase with the sanctions for errors. In contrast, a price cap does not affect the monopoly's selection costs, but it constrains the monopoly's incentives to select by quantity. A fine imposed on non-compliant consumers affects

[49] We derive in Appendix E the conditions such that implementing the imperfect blocking strategy with sanctions for different types of errors is possible and give below an example in which this does not work.

the participation of non-compliant consumers in the market, and thus, the monopoly's optimal response.

An increase in either of regulatory instruments has three effects on social welfare. First, there is a direct cost effect. An increase in the sanctions raises the marginal cost of serving consumers who are not excluded and the marginal cost of missed sales by $\mathbb{E}\left[(1-e)k_\sigma + em_\sigma \,|\mathcal{B}\right]$. This effect is null for the price cap or for the fine. Second, regulatory instruments have an indirect effect on the choice of the selection technology, which implies a variation in the marginal social cost of excluding consumers given by $\mathbb{E}\left[(e'(w-\lambda(r-t)-o)+a')s_\sigma^m \,|\mathcal{B}\right]$. In addition, this variation in the selection technology implies a change in the allocation of consumers between the compliant and non-compliant categories, and their participation in the higher-quality market. Third, the choice of a regulatory instrument has an indirect effect on the monopoly's price, which also implies a change in the allocation of consumers. The choice of the optimal instrument to combine with sanctions for type II errors depends on the resultant of these three effects on social welfare.

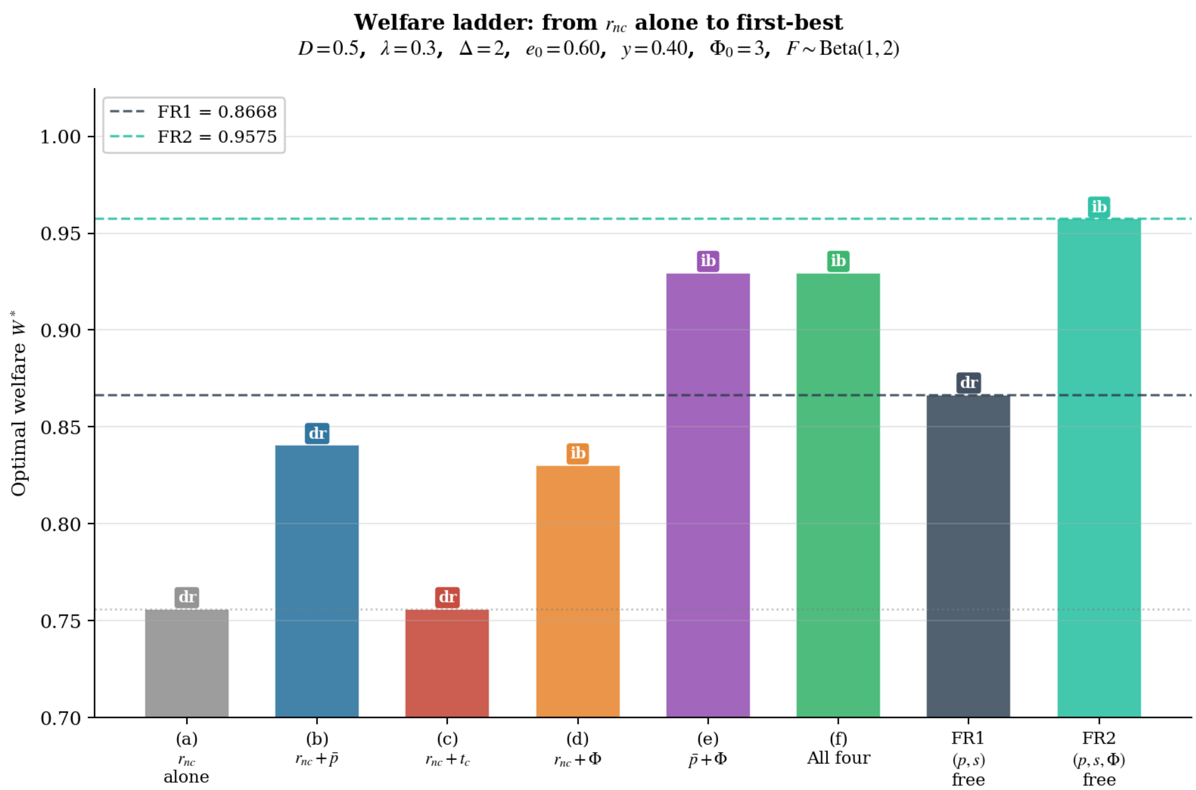
Table 1 compares various instruments that a regulator can use to improve welfare. The line FR2 illustrates a benchmark in which the regulator is able to control s, p and ϕ, the fine imposed on non-compliant consumers. The line FR1 represents the case analyzed in the paper, in which the regulator controls s and p. Cases a) to f) represent second-best situations, in which the regulator uses either r_{nc} alone, a combination of two instruments, or four instruments: r_{nc}, t_c, ϕ, and a price cap $\overline{p}$.

If using a price cap is impossible, we see that the regulator cannot implement the imperfect blocking strategy by only using a sanction for type II errors (r_{nc}). However, combining r_{nc} with a reduction in the fine ϕ imposed on criminals, the regulator succeeds in implementing the imperfect blocking strategy and improves social welfare. This welfare improvement deteriorates the surplus of compliant consumers, who pay a higher price because the willingness to pay of non-compliant consumers increases. In that specific case, adding a sanction for type I errors does not improve the situation because it reduces the monopoly's profit and constrains its participation. Having a price cap could even more improve social welfare than a fine imposed on criminals. The reason is that this instrument increases the participation of compliant consumers and their surplus.

Table 1: comparison of additional regulatory instruments: t_c, $\overline{p}$, and ϕ

Instrument(s)	r_{nc}^{w}	$\bar{p}^{w}$	t_c^{w}	Φ^{w}	W^*	Π^*	CS_C^*	CS_{NC}^*	**Regime**
(a) r_{nc} alone	0.236	—	—	—	0.756	0.065	0.058	0.000	dr
(b) $r_{nc}+\bar{p}$	0.001	0.060	—	—	0.841	0.014	0.193	0.000	dr
(c) $r_{nc}+t_c$	0.236	—	0.000	—	0.756	0.065	0.058	0.000	dr
(d) $r_{nc}+\Phi$	0.001	—	—	0.853	0.830	0.133	0.044	0.020	ib
(e) $\bar{p}+\Phi$	0.001	0.060	—	1.129	0.930	0.023	0.187	0.085	ib
(f) All four	**0.001**	**0.060**	**0.000**	**1.129**	**0.930**	**0.023**	**0.187**	**0.085**	**ib**
FR1: (p,s) free, $\Phi=\Phi_0$	—	—	—	—	0.867	-0.004	0.237	0.000	dr
FR2: (p,s,Φ) free	—	—	—	1.083	0.957	-0.005	0.233	0.096	ib

Figure 3: comparison of additional regulatory instruments: t_c, $\overline{p}$, and ϕ



5.2.3 The surplus of compliant consumers

A policy intervention that aims to influence the monopoly's incentives to exclude non-compliant consumers affects the surplus obtained by compliant ones.[50] From the results of section 4.3, imposing sanctions for mis-classification errors may also reduce the surplus that compliant consumers obtain when they participate in the higher quality market. For example, if $e_c^{'}(s^m) > 0$, a higher sanction for type II errors reduces the surplus that compliant consumers obtain if they buy the monopoly's service rather than the outside option if and only if $db_c/dr_{nc} < 0$, which, from (11), happens if and only

$$-e_c^{'}(s^m)(\Delta y - p^m)s^m_{r_{nc}} < (1 - e_c(s^m))p^m_{r_{nc}}.$$

If $\tilde{\pi}^{ib}_{sp} < 0$, we showed that we have $p^{ib}_{r_{nc}} \geq 0$ and $s^{ib}_{r_{nc}} \leq 0$. Therefore, a higher sanction for type II errors reduces the surplus of compliant consumers if the probability that the monopoly excludes them is not very sensitive to the performance of the algorithm. In this case, consumers have lower incentives to participate in the higher-quality market and comply because they pay a higher price. However, a sanction for type II errors may also increase the surplus of compliant consumers. If the probability that the monopoly excludes them is very sensitive to the sanction and if the sanction does not result in strong price variations, their surplus may increase.

If $\tilde{\pi}^{ib}_{sp} > 0$, we showed that a higher sanction for type II errors sometimes provides the monopoly with incentives to increase its price and the quality of the selection technology; that is, we have $p^{ib}_{r_{nc}} \geq 0$ and $s^{ib}_{r_{nc}} \geq 0$. In this case, a higher sanction for type II errors unambiguously reduces the surplus that compliant consumers obtain if they participate in the higher-quality market.

6 Conclusion

In this paper, we analyzed why a monopoly chooses a quality for a selection technology that differs from the social optimum. Veiga and Weyl (2016) identified three distortions when the monopoly chooses the quality of a product in a selection market: a direct cost effect, a sorting effect, and the Spence distortion. We showed that an additional distortion arises when a monopoly

[50]The same kind of logic applies in platform markets.

chooses the quality of a selection technology, which is caused by the inefficient decision to exclude some consumers from the market. The monopoly's strategy results in an inefficient allocation of consumers between the compliant and non-compliant categories, respectively, and an inefficient total participation in the higher-quality market.

If it is not socially optimal to completely exclude non-compliant consumers from the higher-quality market, we showed that the regulator may need two instruments to implement an imperfect blocking strategy. Because the choice of the selection technology creates an externality between compliant and non-compliant consumers, respectively, having only a sanction for false negatives may have the perverse effect of reducing the surplus of compliant consumers. The overall impact of sanctions on welfare depends on the relative responsiveness of each category of consumers to each type of sanction.

Therefore, AI regulations are likely to have unintended consequences that deserve to be better understood. In our paper, we showed that the choice of the performance of a selection algorithm is likely to create externalities between different consumer categories that a monopoly fails to internalize. We also showed that the choice of the appropriate instruments to correct these external effects is not obvious and depends on a careful analysis of consumer demand response to the choice of the technology. Our results have important implications for the empirical studies conducted on algorithms in several other sectors of the economy. For example, firms may choose the performance of a selection algorithm in credit markets without internalizing the effect of their decisions on credit rationing or interest rates.

References

Acemoglu, D. and Lensman, T. (2024). Regulating transformative technologies. American Economic Review: Insights, 6(3):359–376.

Acemoglu, D., Makhdoumi, A., Malekian, A., and Ozdaglar, A. (2022). Too much data: Prices and inefficiencies in data markets. American Economic Journal: Microeconomics, 14(4):218–256.

Agrawal, A., Gans, J., and Goldfarb, A. (2018). Prediction, judgment and complexity: A theory of decision making and artificial intelligence. NBER Working Paper no 24243.

Akerlof, G. A. (1970). 4. the market for 'lemons': quality uncertainty and the market mechanism. Market Failure or Success, 66.

Battiston, P., Gamba, S., and Santoro, A. (2024). Machine learning and the optimization of prediction-based policies. Technological Forecasting and Social Change, 199:123080.

Becker, G. S. (1968). Crime and punishment: An economic approach. Journal of Political Economy, 76(2):169–217.

Bergemann, D. and Bonatti, A. (2019). Markets for information: An introduction. Annu. Rev. Econ, 11:1–23.

Besanko, D., Donnenfeld, S., and White, L. (1987). Monopoly and quality distortion: Effects and remedies. Quarterly Journal of Economics, 102(4):743–768.

Besfamille, M., Figueroa, N., and Guzmán-Lizardo, L. (2025). Ramsey pricing revisited: Natural monopoly regulation with evaders. The Journal of Industrial Economics, 73(4):569–588.

Biancini, S. and Verdier, M. (2023). Bank-platform competition in the credit market. International Journal of Industrial Organization, 91:103029.

Chen, Y. and Hua, X. (2026). Product safety in the age of ai: Autonomy, r&d, and liability. The Economic Journal.

Colliard, J.-E. (2019). Strategic selection of risk models and bank capital regulation. Management Science, 65(6).

Commission, U. C. (2024). "new data suggests nearly half of charities experience issues when banking". `https://www.gov.uk/government/news/new-data-suggests-nearly-half-of-charities-experience-issues-when-banking`. Accessed on : April 2, 2026.

Cowgill, B., Davis, J. M. V., Montagnes, B. P., and Perkowski, P. (2025). Stable matching on the job? theory and evidence on internal talent markets. Management Science, 71(3):2508–2526.

Creti, A. and Verdier, M. (2014). Fraud, investments and liability regimes in payment platforms. International Journal of Industrial Organization, 35:84–93.

Damiano, E. and Li, H. (2007). Price discrimination and efficient matching. Economic Theory, 30(2):243–263.

Daugherty, P. R. and Wilson, H. J. (2018). Human+ machine: Reimagining work in the age of AI. Harvard Business Press.

Dawid, H. and Muehlheusser, G. (2022). Smart products: Liability, investments in product safety, and the timing of market introduction. Journal of Economic Dynamics and Control, 134:104288.

Einav, L. and Finkelstein, A. (2011). Selection in insurance markets: Theory and empirics in pictures. Journal of Economic Perspectives, 25(1):115–138.

Einav, L., Finkelstein, A., and Cullen, M. R. (2010). Estimating welfare in insurance markets using variation in prices. The Quarterly Journal of Economics, 125(3):877–921.

Farboodi, M. and Veldkamp, L. (2020). Long-run growth of financial data technology. American Economic Review, 110(8):2485–2523.

Fluet, C. and Mungan, M. C. (2025). Oriented data-generating processes: A categorization of roc curves. Theorie and Decision.

Fraisse, H. and Laporte, M. (2021). Return on investment on ai: The case of capital requirement. Banque de France Working Paper no 809.

Gans, J. S. (2023). Artificial intelligence adoption in a monopoly market. Managerial and Decision Economics, 44(2):1098–1106.

Garoupa, N. (1997). The theory of optimal law enforcement. Journal of Economic Surveys, 11(3):267–295.

Garratt, R. J. and Van Oordt, M. R. (2021). Privacy as a public good: a case for electronic cash. Journal of Political Economy, 129(7):2157–2180.

Goh, R. Y. and Lee, L. S. (2019). Credit scoring: A review on support vector machines and metaheuristic approaches. Advances in Operations Research, (1):1974794.

Gurkan, H. and de Véricourt, F. (2022). Contracting, pricing, and data collection under the ai flywheel effect. Management Science, 68(12):8791–8808.

Hua, X. and Spier, K. E. (2023). Platform safety: Strict liability versus negligence. HKUST Bus. Sch. Rsch. Paper no. 2023-105, Mar. 31, 2023.

Hua, X. and Spier, K. E. (2025). Holding platforms liable. American Economic Journal: Microeconomics, 17(4):68–101.

Hurlin, C., Pérignon, C., and Saurin, S. (2024). The fairness of credit scoring models. Management Science, 72(1):406–425.

Ichihashi, S. (2021). Competing data intermediaries. The RAND Journal of Economics, 52(3):515–537.

Jarle Kind, H., Koethenbuerger, M., and Schjelderup, G. (2008). Efficiency enhancing taxation in two-sided markets. Journal of Public Economics, 92(5-6):1531–1539.

Jones, C. I. and Tonetti, C. (2020). Nonrivalry and the economics of data. American Economic Review, 110(9):2819–2858.

Jullien, B., Lefouili, Y., and Riordan, M. H. (2020). Privacy protection, security, and consumer retention. CEPR DP no 15072.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., and Mullainathan, S. (2017). Human decisions and machine predictions. The Quarterly Journal of Economics, 133(1):237–293.

Llanes, G. and Madio, L. (2024). Business strategy and regulation of generative ai firms. Technical report.

Maćkowiak, B., Matějka, F., and Wiederholt, M. (2023). Rational inattention: A review. Journal of Economic Literature, 61(1):226–273.

Mahoney, N. and Weyl, E. G. (2017). Imperfect competition in selection markets. Review of Economics and Statistics, 99(4):637–651.

Markovich, S. and Yehezkel, Y. (2021). "For the public benefit": who should control our data? NET Institute Working Paper Working Paper no 21-08.

Mullainathan, S. and Spiess, J. (2017). Machine learning: An applied econometric approach. Journal of Economic Perspectives, 31(2):87–106.

Mungan, M. C., Obidzinski, M., and Oytana, Y. (2023). Accuracy and preferences for legal error. American Law and Economics Review, 25(1):190–227.

Mussa, M. and Rosen, S. (1978). Monopoly and product quality. Journal of Economic Theory, 18(2):301–317.

Obidzinski, M. and Oytana, Y. (2024). Artificial intelligence, inattention and liability rules. International Review of Law and Economics, 79:106211.

O'Brien, D. P. and Smith, D. (2014). Privacy in online markets: A welfare analysis of demand rotations. TC Bureau of Economics Working Paper no 323.

Polinsky, A. M. and Shavell, S. (2007). The theory of public enforcement of law. Handbook of Law and Economics, 1:403–454.

Rainio, O., Teuho, J., and Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. Scientific Reports, 14(6086).

Rambachan, A., Kleinberg, J., Mullainathan, S., and Ludwig, J. (2020). An economic approach to regulating algorithms. National Bureau of Economic Research Working Paper no 27111.

Rochet, J.-C. and Tirole, J. (2003). Platform competition in two-sided markets. Journal of the european economic association, 1(4):990–1029.

Rochet, J.-C. and Tirole, J. (2006). Two-sided markets : A progress report. the RAND Journal of Economics, 35(3):645–667.

Sarlin, P. (2013). On policymakers' loss functions and the evaluation of early warning systems. Economics Letter, 119(1).

Segal, I. (1999). Contracting with externalities. The Quarterly Journal of Economics, 114(2):337–388.

Sims, C. A. (2003). Implications of rational inattention. Journal of Monetary Economics, 50(3):665–690.

Spence, A. M. (1975). Monopoly, quality, and regulation. The Bell Journal of Economics, pages 417–429.

Spier, K. E. and Van Loo, R. (2025). Foundations for plateform liability. Notre-Dame Law Review, 100(3):1137–1200.

Stiglitz, J. E. (1977). Monopoly, non-linear pricing and imperfect information: the insurance market. The Review of Economic Studies, 44(3):407–430.

Takáts, E. (2011). A theory of "crying wolf": The economics of money laundering enforcement. The Journal of Law, Economics, & Organization, 27(1):32–78.

Van Loo, R. (2020). The new gatekeepers. Virginia Law Review, 106(2):467–522.

van Oordt, M. (2025). Transforming the digital euro proposal from a threat into an opportunity for privacy. In Frontiers of Digital Finance, pages 63–72. CEPR Press.

Veiga, A. and Weyl, E. G. (2016). Product design in selection markets. The Quarterly Journal of Economics, 131(2):1007–1056.

Veiga, A., Weyl, E. G., and White, A. (2017). Multidimensional platform design. American Economic Review, 107(5):191–195.

Weyl, E. G. (2010). Price theory of multi-sided platforms. American Economic Review, 100(4).

Zhang, Y. and Trubey, P. (2019). Machine learning and sampling scheme: An empirical study of money laundering detection. Computational Economics, 54(3):1043–1063.

Appendix

Appendix A: Consumer demand for the service and some preliminary results

Appendix A-1: Consumer demand for the service

To determine whether the market is covered or not, we compare the indifferent consumers given in the analysis of consumer demand. From (12) and (14), we have $b_I \leq b_c$ if and only if

$$\big(1+\Delta(1-e_{nc})\big)(b_I - b_c) = -b_c \frac{\big(1+\Delta(1-e_c)\big)(1-e_{nc})}{(1-e_c)} + \Delta(1-e_{nc})b_\phi \leq 0.$$

Moreover, from (12) and (11), we have $b_{nc} \leq b_c$ if and only if $p \leq \underline{p}$, where

$$\underline{p} \equiv \Delta y - \frac{\Delta b_\phi}{1+\Delta(1-e_c)}.$$

In addition, we have $b_I \leq b_{nc}$ if and only if $p \geq \underline{p}$ and $b_{nc} \leq b_c$ if and only if $p \leq \underline{p}$.

In summary, consumer demand for the firm's service depends on the price of the service and the quality of the selection technology:

- If $p > \Delta y$, we have $\mathcal{B}_c = \emptyset$ and $\mathcal{B}_{nc} = (min\{b_{nc}, B\}, B)$.
- If $\underline{p} \leq p \leq \Delta y$, with
$$\underline{p} \equiv \Delta y - \frac{\Delta b_\phi}{1+\Delta(1-e_c)},$$
we have $b_c \leq b_I \leq b_{nc}$. The market is not covered, with $\mathcal{B}_c = (0, min\{b_c, B\})$ and $\mathcal{B}_{nc} = (min\{b_{nc}, B\}, B)$.
- If $p < \underline{p}$, we have $b_{nc} \leq b_I \leq b_c$, and the market is covered, with $\mathcal{B}_c = (0, max\{0, b_I\})$ and $\mathcal{B}_{nc} = (min\{b_I, B\}, B)$.

Since $b_c = (1-e_c)(\Delta y - p)$ and $D_c = F(b_c)$, the demand of compliant consumers decreases with e_c. Since $b_\phi(s) = \frac{\phi e_{nc}(s)}{\Delta(1-e_{nc}(s))}$, $b_{nc} = (p/\Delta) - y + b_\phi(s)$ and $D_{nc} = 1 - F(b_{nc})$, the demand of non-compliant consumers decreases with e_{nc}.

This completes the proof of Lemma 1.

Appendix A-2: Proof of Lemma 2

To succeed in excluding non-compliant consumers, from (12), the monopoly must choose a price p for its service and a quality of the selection technology $s \in [0, S]$ so that $b_{nc} \geq B$. From Eq.(12), this implies that the price p of the service and the quality s of the selection technology should be chosen so that $p \geq \Delta(B + y - b_\phi(s))$ or $s = S$ because with the maximal quality, non-compliant consumers are excluded with certainty. We denote by $\hat{p}(s) \equiv \Delta(B + y - b_\phi(s))$ the minimum price such that non-compliant consumers are completely excluded. Since b_ϕ increases with s from Eq.(13) if $s \in [\bar{s}, S]$, we have that $\hat{p}(s)$ decreases with s for $s \in [\bar{s}, S]$. This completes the proof of Lemma 2.

Appendix A-3: Extending the model of Veiga and Weyl (2016) to our setting

To make our results comparable to their paper, we adapt the definitions given by Veiga and Weyl (2016) to our framework. In our setting, the consumer's type $\Theta = (b, i)$ is bi-dimensional (whereas they consider a finite arbitrary number of dimensions). There are three differences in our setting:

1. One of the dimensions of the consumer's type (i.e., i) is a discrete variable.

2. The value of purchasing the outside option depends on the consumer's type. This implies that b_c and b_{nc} do not have the same monotonicity with respect to price variations (unlike the parameter $\tilde{\tau}$ of their paper, which is strictly increasing with p under their assumptions). In our paper, from (11) and (12), b_c is decreasing with p, whereas b_{nc} is increasing with p.

3. The firm may exclude some consumers from the market, which implies that it considers the expected margin associated with a selection technology. Therefore, the quality of the selection technology has an additional effect on the exclusion of consumers, which depends on their types.

We consider the case in which the market is not covered. Using the implicit function theorem on

Eq.(1) given by $\tilde{u}(s, b_i, i) - (1 - e_i)p - (y + b_i) = 0$ yields:

$$\frac{\partial b_i}{\partial s} = -\frac{(\tilde{u}'(e_i, b) + e_i'(s)p)}{\frac{\partial \tilde{u}(e_i,b)}{\partial b} - 1}, \tag{23}$$

and

$$\frac{\partial b_i}{\partial p} = \frac{1 - e_i}{\frac{\partial \tilde{u}(e_i,b)}{\partial b} - 1}. \tag{24}$$

In our particular setting, we have:

$$\frac{\partial \tilde{u}(e_i, b_i, i)}{\partial b} - 1 = l\Delta(1 - e_i) + l - 1, \tag{25}$$

and

$$\tilde{u}'(e_i, b, i) = -\Big(\Delta(lb + y) + l\phi\Big)e_i'(s),$$

with $l = 1$ if the consumer is not compliant and $l = 0$ otherwise. Therefore, we have

$$\frac{\partial b_c}{\partial p} = -(1 - e_c), \tag{26}$$

$$\frac{\partial b_{nc}}{\partial p} = \frac{1}{\Delta}. \tag{27}$$

$$\frac{\partial b_c}{\partial s} = -(\Delta y - p)e_c^{'}(s), \tag{28}$$

and

$$\frac{\partial b_{nc}}{\partial s} = \frac{\partial b_\phi}{\partial s} = \frac{\phi}{\Delta}\frac{e_{nc}^{'}(s)}{\big(1 - e_{nc}(s)\big)^2}. \tag{29}$$

Since

$$Q(s, p) = \sum_{i \in \{c, nc\}} \int_{\mathcal{B}_i} f(b)db,$$

and from (23), differentiating Q with respect to s gives:

$$\frac{\partial Q}{\partial s} = \frac{\partial b_c}{\partial s}f(b_c) - \frac{\partial b_{nc}}{\partial s}f(b_{nc}) = M\mathbb{E}\left[\tilde{u}' + e'p \,\middle|\, \mathcal{M}\right].$$

Then, since $M = -\partial Q/\partial p$, using the implicit function theorem on the equation $Q\big(s, P(s,q)\big) = q$ that defines the inverse demand function yields:

$$\frac{\partial P}{\partial s} = \mathbb{E}\left[\tilde{u}' + e'p \,|\mathcal{M}\right]. \tag{30}$$

The proof of the results follows Appendix A-1 of Veiga and Weyl (2016). We tackle the maximization of social welfare and profit simultaneously by defining:

$$Z(s,q) = \sum_{i\in\{c,nc\}} \int_{\mathcal{B}_i} z(s,q) f(b) db.$$

Profit maximization considers $z(s,q) = \big(1 - e(s,\theta)\big)\big(P(s,q) - k(s,\theta)\big) - e(s,\theta)m(s,\theta)$, whereas welfare maximization considers $z(s,q) = \big(1 - e(s,\theta)\big)\big(w(\theta) - k(s,\theta) - o(\theta)\big) - e(s,\theta)m(s,\theta)$.

The FOC with respect to s is given by:

$$\sum_{i\in\{c,nc\}} \int_{\mathcal{B}_i} \frac{\partial z(s,q)}{\partial s} f(b) db + \sum_{i\in\{c,nc\}} (-1)^l (\frac{\partial b_i}{\partial s} + \frac{\partial b_i}{\partial p}\frac{\partial P}{\partial s}) z(s,q) = 0,$$

with $l = 0$ if the consumer is compliant and $l = 1$ otherwise. Replacing (23), (24) and (30) in the FOC with respect to s and using the notations (8) and (9) gives:

$$Q\mathbb{E}\left[\frac{\partial z}{\partial s}\middle|\mathcal{B}\right] + M(\mathbb{E}(\tilde{u} + e'p)z\,|\mathcal{M}) - \mathbb{E}((1-e)z\,|\mathcal{M})\,\mathbb{E}((\tilde{u} + e'p)\,|\mathcal{M})) = 0,$$

or else:

$$Q\mathbb{E}\left[\frac{\partial z}{\partial s}\middle|\mathcal{B}\right] + M(\mathrm{Cov}\left[\tilde{u}' + e'p, z\,|\mathcal{M}\right] + \mathbb{E}(ez\,|\mathcal{M})\,\mathbb{E}((\tilde{u} + e'p)\,|\mathcal{M})) = 0.$$

Replacing for $MS = Q/M$ the marginal consumer surplus, after simplification by $Q > 0$, the FOC with respect to s is given by:

$$\mathbb{E}\left[\frac{\partial z}{\partial s}\middle|\mathcal{B}\right] + \frac{1}{MS}(\mathrm{Cov}\left[\tilde{u}' + e'p, z\,|\mathcal{M}\right] + \mathbb{E}(ez\,|\mathcal{M})\,\mathbb{E}((\tilde{u} + e'p)\,|\mathcal{M})) = 0. \tag{31}$$

The FOC with respect to q is given by:

$$\sum_{i\in\{c,nc\}}\int_{\mathcal{B}_i}\frac{\partial z(s,q)}{\partial q}f(b)db+\sum_{i\in\{c,nc\}}(-1)^l(\frac{\partial b_i}{\partial p}\frac{\partial P}{\partial q})z(s,q)=0,$$

with $l=0$ if the consumer is compliant and $l=1$ otherwise and $\partial P/\partial q=-1/M$. Replacing (24) in the FOC with respect to q gives:

$$\sum_{i\in\{c,nc\}}\int_{\mathcal{B}_i}\frac{\partial z(s,q)}{\partial q}f(b)db+\mathbb{E}((1-e)z\,|\mathcal{M})=0. \tag{32}$$

Appendix A-4: Examples of distributions

Table 0: Parameters for the examples

Parameter / Outcome	Config. 1 (decreasing)	Config. 2 (increasing)
Model parameters		
e_0 (common initial value)	0.30	0.40
c (curvature of e_{nc})	1.20	1.50
$s^{enc}_{\min}=\frac{1}{2c}$	0.417	0.333
y (income)	0.60	0.60
Φ (detection intensity)	2.0	2.0
c_c (detection cost, compliant)	0.30	0.20
c_{nc} (detection cost, non-compliant)	0.45	0.30
F (distribution)	Beta(1,2)	Beta(1,2)
Detection functions		
$e_c(s)$	$e_0(1-s)$	$e_0(1-s)$
$e_{nc}(s)$	$e_0(1-s+1.2s^2)$	$e_0(1-s+1.5s^2)$
$e_c(0)=e_{nc}(0)$	0.30 ✓	0.40 ✓
$e_{nc}(s)-e_c(s)=e_0cs^2$	increasing ✓	increasing ✓
Both-types regime results		
Monotonicity of $s^{ib}(r_{nc})$	Decreasing	Increasing
Range of s^{ib}	[0.283, 0.336]	[0.372, 0.431]
Amplitude of s^{ib}	0.0532	0.0585
Monotonicity of $p^{ib}(r_{nc})$	Increasing	Increasing
Range of p^{ib}	[0.765, 0.882]	[0.717, 0.792]
Shared parameters		
Δ (surplus gain)	2.0	2.0
B (distribution support)	1.0	1.0
Mechanism (ds^{ib}/dr_{nc})	$s^*<s^{enc}_{\min}$, $e'_{nc}(s^*)<0$	$s^*>s^{enc}_{\min}$, $e'_{nc}(s^*)>0$

Notes: $e_c(s)=e_0(1-s)$, $e_{nc}(s)=e_0(1-s+cs^2)$. *Both configs:* $e_c(0)=e_{nc}(0)=e_0$; $e_{nc}(s)-e_c(s)=e_0cs^2>0$; e_{nc} *U-shaped with minimum at* $s^{enc}_{\min}=\frac{1}{2c}$. *Red = decreasing, Blue = increasing.*

Appendix B: Profit maximization

Appendix B-1: Profit maximization with the imperfect blocking strategy

Profit maximization with respect to s:

We apply the result of Appendix A-3 given in (31) to the function $z = \pi$ with

$$\pi(s,q) = \big(1 - e(s,\theta)\big)\big(P(s,q) - k(s,\theta)\big) - e(s,\theta)m(s,\theta).$$

Omitting (s,θ) when possible, when the firm maximizes its profit with respect to s, we have:

$$\frac{\partial z}{\partial s} = \big(1-e\big)\frac{\partial P}{\partial s} - \big(1-e\big)k' - em' - \big(P(s,q) - k + m\big)e',$$

with $\partial P/\partial s = \mathbb{E}\left[\tilde{u}' + e'p\,|\mathcal{M}\right]$ given in (30). If there is an interior solution, from (31), the FOC of the firm's profit-maximization with respect to s is given by:

$$\begin{aligned}&-\mathbb{E}\left[(1-e)k' + em'\,\big|\mathcal{B}\right] + \mathbb{E}\left[\tilde{u}' + e'p\,\big|\mathcal{M}\right]\mathbb{E}\left[1-e\,|\mathcal{B}\right] - \mathbb{E}\left[e'(P-k+m)\,\big|\mathcal{B}\right]\\ &+\frac{\mathrm{Cov}\left[\tilde{u}' + e'p, \pi\,|\mathcal{M}\right] + \mathbb{E}(e\pi\,|\mathcal{M})\,\mathbb{E}((\tilde{u} + e'p)\,|\mathcal{M})}{MS} = 0.\end{aligned}$$

Profit maximization with respect to p or q:

We apply the result of Appendix A-3 given in (32) to the function $z = \pi$. Since $\partial P/\partial q = -1/M$, we have $\partial \pi/\partial q = -(1-e)/M$. After multiplication by $M > 0$, we obtain:

$$M\mathbb{E}\left[(1-e)\pi\,\big|\mathcal{M}\right] - \mathbb{E}\left[(1-e)\,|\mathcal{B}\right] = 0.$$

A special case with $e = 0$:

If $e = 0$, we have $\mathbb{E}\left[(1-e)\,|\mathcal{B}\right] = Q$ and $\mathbb{E}\left[(1-e)^2(p-k) - e(1-e)m\,\big|\mathcal{M}\right] = p - \mathbb{E}\left[k\,|\mathcal{M}\right]$. Therefore, after simplification by $M > 0$, we find the same result as Veiga and Weyl (2016), that is, the FOC of profit maximization with respect to q (or p) is given by

$$p - \mathbb{E}\left[k \,|\mathcal{M}\right] - \frac{Q}{M} = 0.$$

Conditions such that there is an interior solution:

From Lemma 3, for an interior solution to exist, it must be that $p^{ib} > \underline{p}(s^{ib})$ and $p^{ib} < \hat{p}(s^{ib})$. If $p^{ib} \geq \hat{p}(s^{ib})$, the monopoly prefers the de-risking strategy. If $p^{ib} \leq \underline{p}(s^{ib})$, the market is covered.

Second-order conditions of profit-maximization (existence of an interior solution):

We denote by $\alpha = \frac{\partial^2 \pi^{ib}}{\partial^2 p}\Big|_{(p^{ib},s^{ib})}$, $\beta = \frac{\partial^2 \pi^{ib}}{\partial^2 s}\Big|_{(p^{ib},s^{ib})}$, and $\gamma = \frac{\partial^2 \pi^{ib}}{\partial p \partial s}\Big|_{(p^{ib},s^{ib})}$.

For a local interior maximum to exist at (p^{ib}, s^{ib}), it must be that $\alpha < 0$ and $\delta = \alpha\beta - \gamma^2 > 0$.

The second derivative of π^{ib} with respect to p is given by:

$$\alpha = \frac{\partial^2 \pi^{ib}}{\partial^2 p}\Big|_{(p^{ib},s^{ib})} = -2(1-e_c)^2 f(b_c) - \frac{2}{\Delta}(1-e_{nc})f(b_{nc}) + \tilde{\pi}^c(1-e_c)^2 f'(b_c) - \tilde{\pi}^{nc} f'(b_{nc})(1/\Delta).$$

For example, if f is the density of a uniform distribution with $f' = 0$, we have $\alpha < 0$.

The second derivative of π^{ib} with respect to s is given by:

$$\beta = \frac{\partial^2 \pi^{ib}}{\partial^2 s}\Big|_{(p^{ib},s^{ib})} = (\tilde{\pi}^c_s)' D_c + (\tilde{\pi}^{nc}_s)' D_{nc} + 2\tilde{\pi}^c_s D^{'}_c + 2\tilde{\pi}^{nc}_s D^{'}_{nc}.$$

Since $\alpha < 0$, for $\alpha\beta - \gamma^2$ to be strictly positive, it must be that $\beta < 0$. Since $D^{'}_{nc} < 0$, $\tilde{\pi}^{nc}_s = -e^{'}_{nc}(p - r_{nc}) - a^{'}_{nc} < 0$, we have $\tilde{\pi}^{nc}_s D^{'}_{nc} > 0$. Therefore, it must be that at (p^{ib}, s^{ib}), the monopoly's profit from each category of consumers is sufficiently concave with respect to s. This condition is necessary but not sufficient. However, we see for instance, that β is strictly negative if a_{nc} or e_{nc} are sufficiently convex with respect to s.

The cross derivative of π^{ib} with respect to s and p is given by:

$$\gamma = \frac{\partial^2 \pi^{ib}}{\partial s \partial p}\Big|_{(p^{ib},s^{ib})} = -e^{'}_c D_c - \tilde{\pi}^c_s(1-e_c)f(b_c) + \tilde{\pi}^c(e^{'}_c f(b_c) + e^{'}_c b_c f'(b_c)) - e^{'}_c b_c f(b_c)$$

$$-e_{nc}^{'}D_{nc} - (1-e_{nc})b_{\phi}^{'}f(b_{nc}) - \tilde{\pi}_s^{nc}(1/\Delta)f(b_{nc}) - \tilde{\pi}^{nc}(1/\Delta)b_{\phi}^{'}f'(b_{nc}),$$

with $e_{nc}^{'} > 0$, $b_{\phi}^{'} > 0$, $\tilde{\pi}_s^{nc} < 0$, $\tilde{\pi}_s^{c} = -(e_c^{'})(p+t) - a_c^{'}$. It is important to note that this cross derivative may be positive or negative. For example, if $e_c^{'}$ is close to zero, and if the monopoly does not lose marginally to much from the non-compliant consumers when s increases (i.e, $\tilde{\pi}^{nc}$ small), the coefficient γ is negative. For example, if b follows a uniform distribution on $(0,1)$, $t_c = a_c = a_{nc} = 0$ and e_{nc} is a constant, we have:

$$\gamma = 2e_c^{'}(1-e_c)(2p^{ib} - \Delta y),$$

which is positive if $e_c^{'} > 0$ because $p^{ib} \geq (\Delta y)/2$, which is the price that the monopoly chooses if it only sells to compliant consumers. If $F(0) > 0$, if e_{nc} is a constant ($e_{nc}^{'} = 0$), if e_c becomes close to 1, b_c is close to zero, and we see that

$$\gamma = -e_c^{'}F(0),$$

which is negative if $e_c^{'} > 0$.

Appendix B-2: Special case with $a_c(s) = a_{nc}(s) = 0$ and $t_c = t_{nc} = 0$

If there is an interior solution, the first-order condition of profit-maximization with respect to p is given by:

$$\sum_{i \in \{c,nc\}} (1-e_i)(\frac{\partial D_i}{\partial p}(p - k_i(s,r)) + D_i(s,p)) = 0.$$

Introducing the price elasticities of consumer demands gives

$$\sum_{i \in \{c,nc\}} \tilde{D}_i \epsilon_i (\frac{p^{ib} - k_i}{p^{ib}} - \frac{1}{\epsilon_i}) = 0.$$

Since $k^{'} = 0$, introducing $\tilde{\eta}_i = \frac{-\partial(1-e_i)D_i}{\partial s}\frac{s}{(1-e_i)D_i}$, the first-order condition of profit-maximization with respect to s is given by:

$$-(p-k_c)\frac{\tilde{D}_c\tilde{\eta}_c}{s} - (p-k_{nc})\frac{\tilde{D}_{nc}\tilde{\eta}_{nc}}{s} = 0.$$

Multiplying by $s > 0$ gives:

$$-(p - k_c)\tilde{D}_c\tilde{\eta}_c - (p - k_{nc})\tilde{D}_{nc}\tilde{\eta}_{nc} = 0,$$

or else:

$$\frac{\tilde{D}_c(p^{ib} - k_c)}{\tilde{D}_{nc}(p^{ib} - k_{nc})} = -\frac{\tilde{\eta}_{nc}}{\tilde{\eta}_c}$$

For an interior solution to exist so that the monopoly serves both categories of consumers, there must exist a price p and a quality s for the selection technology such that $\tilde{\eta}_c$ and $\tilde{\eta}_{nc}$ have opposite signs. By assumption, we have that $\partial((1 - e_{nc})D_{nc})/\partial s < 0$. Therefore, we have that $\tilde{\eta}_{nc} \geq 0$. Therefore, a necessary condition is that there exist a price p and a quality s for the selection technology such that $\tilde{\eta}_c \leq 0$.

Appendix B-3: Proposition 3: Profit maximization with the de-risking strategy

If the constraint $p > \hat{p}(s)$ is binding, the de-risking profit is given by:

$$\pi = (1 - e_c(s))D_c(s, \hat{p}(s))(\hat{p}(s) - k_c(s)) - e_c(s)m_c(s)D_c(s, \hat{p}(s)),$$

with $\hat{p}(s) = \Delta\big(B + y - b_\phi(s)\big)$, $r_{nc} = 0$, $k_c(s) = a_c(s)$ and $m_c(s) = t_c$.

For this strategy to be possible, two conditions must be verified:

- (1) the exclusion of non-compliant consumers is possible with a price that exceeds the marginal cost $a_c(s)$, that is, there exists $s \in [0, S]$ such that $\hat{p}(s) \geq a_c(s)$, or else $b_\phi(s) \leq (B + y) - a_c(s)/\Delta$,
- (2) the exclusion of non-compliant consumers does not imply the complete exclusion of compliant consumers, that is, we have $b_c(s, \hat{p}(s)) \geq 0$. This implies that there exists $s \in [0, S]$ such that $b_\phi(s) \geq B$.

Replacing for $b_\phi(s)$ given in Eq. (13), combination of these two conditions requires that there exists

$s \in [0, S]$ such that:

$$B \leq \frac{\phi e_{nc}(s)}{\Delta(1 - e_{nc}(s))} \leq (B + y) - a_c(s)/\Delta,$$

which implies after multiplication by $\Delta/\phi > 0$ that

$$\frac{\Delta B}{\phi} \leq \frac{e_{nc}(s)}{(1 - e_{nc}(s))} \leq \frac{\Delta(B + y) - a_c(s)}{\phi},$$

or else:

$$\frac{\phi}{\Delta(B + y) - a_c(s)} + 1 \leq \frac{1}{e_{nc}(s)} \leq \frac{\phi}{\Delta B} + 1.$$

Therefore, the combination of condition (1) and (2) requires that there exists $s \in [0, S]$ such that

$$\frac{\Delta B}{\phi + \Delta B} \leq e_{nc}(s) \leq \frac{\Delta(B + y) - a_c(s)}{\phi + \Delta(B + y) - a_c(s)}. \tag{33}$$

Since e_{nc} and a_c are increasing with s, the function

$$e_{nc}(s) - \frac{\Delta(B + y) - a_c(s)}{\phi + \Delta(B + y) - a_c(s)}$$

is increasing with s. At $s = 0$, since $a_c(0) = 1$ and $e_{nc}(0) = e_0$, it is equal to

$$e_0 - \frac{\Delta(B + y)}{\phi + \Delta(B + y)},$$

which must be negative so that there exists $s \in [0, S]$ satisfying to condition (33). This inequality represents assumption (A3).

The total derivative of π with respect to s is given by

$$\frac{d\pi}{ds} = \frac{\partial \pi}{\partial p} \frac{d\hat{p}(s)}{ds} + \frac{\partial \pi}{\partial s}. \tag{34}$$

Since $\hat{p}(s) = \Delta(B + y - b_\phi(s))$ and $\Delta \partial b_\phi / \partial s = \Delta \partial b_{nc} / \partial s = -B\Delta D'_{nc}$, we have

$$d\hat{p}(s)/ds = B\Delta D'_{nc}. \tag{35}$$

Since $\eta_{nc} = \frac{-\partial D_{nc}}{\partial s} \frac{s}{D_{nc}}$, we have $\frac{\partial D_{nc}}{\partial s} = -\frac{\eta_{nc} D_{nc}}{s}$. We also have

$$\frac{\partial \pi}{\partial p} = (1 - e_c(s))(M_c(p^{dr} - k_c) + D_c), \tag{36}$$

Replacing (35) and (35) into (34) gives

$$\frac{d\pi}{ds} = B\Delta D'_{nc}((1 - e_c(s))M_c(p^{dr} - k_c) + \tilde{D}_c) + \frac{\partial \pi}{\partial s}.$$

If there is an interior solution, the first-order condition of profit-maximization is given by:

$$\hat{p}(s^{dr}) - k_c = -\frac{D_c}{M_c} - \frac{\frac{\partial \pi}{\partial s}}{B\Delta D'_{nc}},$$

with

$$\frac{\partial \pi}{\partial s} = \tilde{D}'_c(\hat{p}(s^{dr}) - k_c) - k'_c \tilde{D}_c - e'_c m_c D_c - e_c m_c D'_c.$$

After a division by $\hat{p}(s^{dr}) > 0$, this gives

$$\frac{\hat{p}(s^{dr}) - k_c}{\hat{p}(s^{dr})} = \frac{1}{\epsilon_c} - \frac{\frac{\partial \pi}{\partial s}}{\hat{p}(s^{dr}) B\Delta D'_{nc}}.$$

If $a_c = k_c = 0$ and $t_c = 0$, we have $k'_c = 0$ and $m_c = 0$, which implies that

$$\frac{\partial \pi}{\partial s} = \tilde{D}'_c \hat{p}(s^{dr}).$$

If at $s = s^{dr}$ we have

$$\Delta b_\phi(s^{dr}) + a_c(s^{dr}) > \Delta(B + y),$$

the monopoly is constrained to sell at its marginal cost and makes zero profit.

This completes the proof of Proposition 3.

Appendix B-4: Comparative statics

Suppose that there is no cost of data collection, that is, $a = 0$. Using the implicit function theorem on each of the first-order conditions of profit-maximization with respect to s and p gives

$$\begin{pmatrix} \frac{\partial p^{ib}}{\partial r_{nc}} \\ \frac{\partial s^{ib}}{\partial r_{nc}} \end{pmatrix} = \frac{1}{\delta} \begin{pmatrix} -\beta(\frac{\partial^2 \pi^{ib}}{\partial p \partial r_{nc}}) + \gamma(\frac{\partial^2 \pi^{ib}}{\partial s \partial r_{nc}}) \\ \gamma(\frac{\partial^2 \pi^{ib}}{\partial p \partial r_{nc}}) - \alpha(\frac{\partial^2 \pi^{ib}}{\partial s \partial r_{nc}}) \end{pmatrix},$$

and

$$\begin{pmatrix} \frac{\partial p^{ib}}{\partial t_c} \\ \frac{\partial s^{ib}}{\partial t_c} \end{pmatrix} = \frac{1}{\delta} \begin{pmatrix} -\beta(\frac{\partial^2 \pi^{ib}}{\partial p \partial t_c}) + \gamma(\frac{\partial^2 \pi^{ib}}{\partial s \partial t_c}) \\ \gamma(\frac{\partial^2 \pi^{ib}}{\partial p \partial t_c}) - \alpha(\frac{\partial^2 \pi^{ib}}{\partial s \partial t_c}) \end{pmatrix},$$

with $\alpha < 0$, $\beta < 0$,

$$\frac{\partial^2 \pi^{ib}}{\partial p \partial r_{nc}} = (1 - e_{nc}) f(b_{nc})/\Delta > 0,$$

$$\frac{\partial^2 \pi^{ib}}{\partial p \partial t_c} = e_c(1 - e_c) f(b_c) > 0,$$

$$\frac{\partial^2 \pi^{ib}}{\partial s \partial r_{nc}} = D^{'}_{nc} e^{'}_{nc} < 0,$$

and

$$\frac{\partial^2 \pi^{ib}}{\partial s \partial t_c} = -e_c D^{'}_c - e^{'}_c D_c.$$

Although this occurs more rarely in practice, the regulator can also decide to impose a sanction for type I errors, when the monopoly excludes compliant consumers.[51] Similarly, we denote by $p^{ib}_{t_c}$ and $s^{ib}_{t_c}$ the marginal impact of an increase in t_c the sanction for type I errors on the price of the service and the quality of the selection technology, respectively. Applying the implicit function theorem to the first-order conditions of profit-maximization yields:

$$p^{ib}_{t_c} = \frac{1}{\Delta\delta}(e_c(1 - e_c) f(b_c) \tilde{\pi}^{ib}_s + (-e_c D^{'}_c - e^{'}_c D_c) \tilde{\pi}^{ib}_{sp}), \tag{37}$$

[51] For example, in the Societe Generale case, the ACPR, the French prudential authority, sanctioned the fact that, for customers benefiting from a legal right of access to a bank account, the bank failed to properly implement its operational obligations: incomplete basic services, irregular account closure and inadequate controls. The bank had to pay a financial penalty of 800,000 euros. The reasoning is that banks using risk-based approaches to select their consumers should not be allowed to exclude entire categories of consumers from the access to basic services.

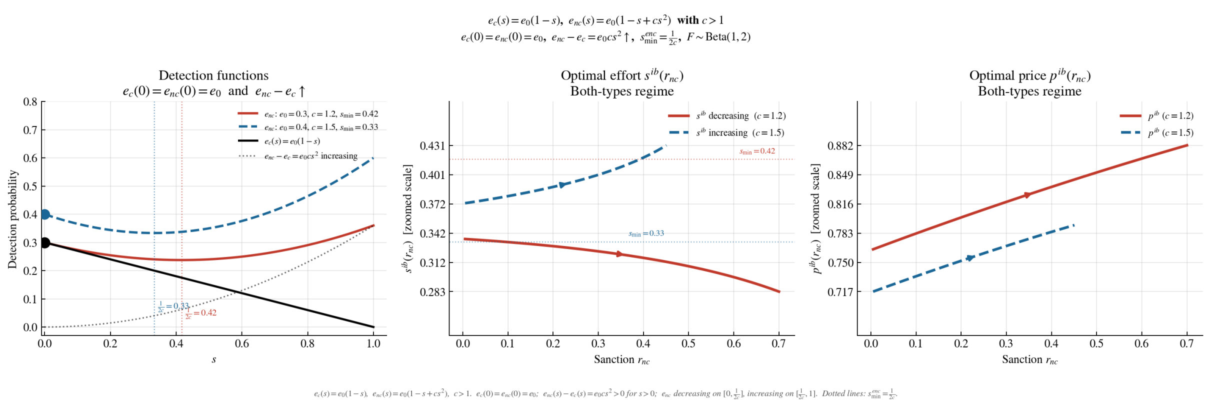
and

$$s_{tc}^{ib} = \frac{1}{\Delta\delta}((-e_c D_c^{'} - e_c^{'} D_c)\tilde{\pi}_p^{ib} + e_c(1-e_c)f(b_c)\tilde{\pi}_{sp}^{ib}), \tag{38}$$

where $e_c^{'} \geq 0$, $\delta > 0$, $\tilde{\pi}_p^{ib} < 0$, $\tilde{\pi}_s^{ib} < 0$ and $-e_c D_c^{'} - e_c^{'} D_c < 0$ if and only if the expected volume of missed sales $e_c D_c$ decreases with s.

Suppose that the expected volume of missed sales decreases with s. If $\tilde{\pi}_{sp}^{ib} < 0$, a higher sanction for type I errors has an ambiguous impact on the price of the service and the quality of the selection technology. If $\tilde{\pi}_{sp}^{ib} > 0$, a higher sanction for type I errors increases the quality of the selection technology and decreases the price of the service; that is, we have $p_{tc}^{ib} \leq 0$ and $s_{tc}^{ib} \geq 0$. Therefore, a higher sanction for type I errors implies that the monopoly has higher incentives to select by technology and lower its price. In contrast, if the expected volume of missed sales increases with s and $\tilde{\pi}_{sp}^{ib} < 0$, a higher sanction for type I errors lowers the price of the service and increases the quality of the selection technology; that is, we have $p_{tc}^{ib} \leq 0$ and $s_{tc}^{ib} \geq 0$.

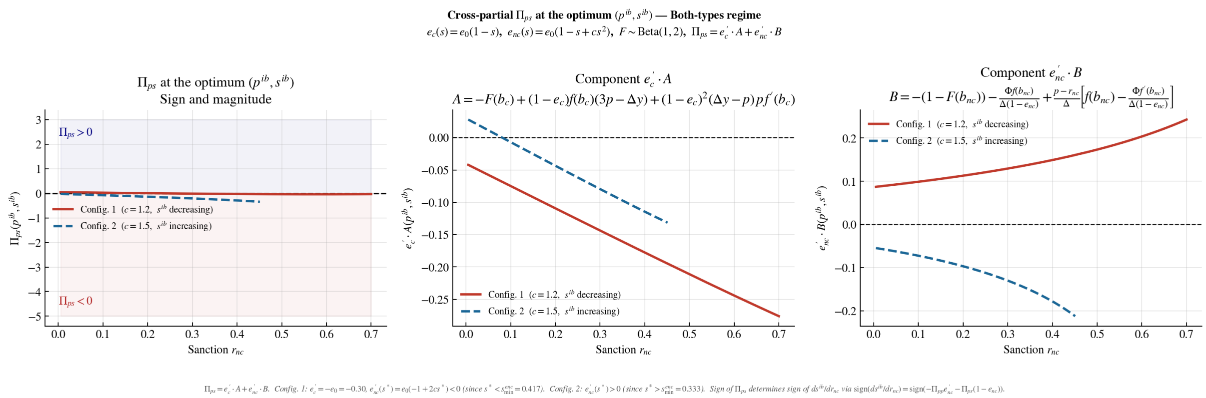
Figure 4: variations of s^{ib} with r_{nc}



Similarly, applying the implicit function theorem on each of the first-order conditions of profit-maximization with respect to s and p gives

$$\begin{pmatrix} \frac{\partial p^{ib}}{\partial \phi} \\ \frac{\partial s^{ib}}{\partial \phi} \end{pmatrix} = \frac{1}{\delta} \begin{pmatrix} -\beta(\frac{\partial^2 \pi^{ib}}{\partial p \partial \phi}) + \gamma(\frac{\partial^2 \pi^{ib}}{\partial s \partial \phi}) \\ \gamma(\frac{\partial^2 \pi^{ib}}{\partial p \partial \phi}) - \alpha(\frac{\partial^2 \pi^{ib}}{\partial s \partial \phi}) \end{pmatrix}$$

Figure 5: sign of the cross-derivative at the optimum (p^{ib}, s^{ib})



with $\alpha < 0$, $\beta < 0$,

$$\frac{\partial^2 \pi^{ib}}{\partial p \partial \phi} = -f(b_{nc})e_{nc}/\Delta - e_{nc}f'(b_{nc})(p - r_{nc})/\Delta,$$

and

$$\frac{\partial^2 \pi^{ib}}{\partial s \partial \phi} = \frac{-(p - r_{nc})e_{nc}^{'}(s)}{\Delta}\left(f(b_{nc}) + \frac{\phi e_{nc} f'(b_{nc})}{\Delta(1 - e_{nc})^2}\right).$$

Suppose that the distribution f is uniform and $e_{nc}^{'}(s) > 0$. If $\gamma > 0$, the price p^{ib} decreases with the fine ϕ imposed on non-compliant consumers. Otherwise, if $\gamma < 0$, the price p^{ib} is non monotonic with the fine ϕ imposed on non-compliant consumers. If $\gamma < 0$, the quality of the selection technology s^{ib} increases with the fine ϕ imposed on non-compliant consumers. Otherwise, if $\gamma > 0$, the quality of the selection technology s^{ib} is non monotonic with the fine ϕ imposed on non-compliant consumers.

Appendix C: Welfare maximization

Appendix C-1: Welfare maximization with respect to s and p:

Welfare maximization with respect to s:

We apply the result of Appendix A-3 given in (31) to the function $z = \tilde{w}$ with

$$\tilde{w}(s, q) = \big(1 - e\big)\big(w - k - o\big) - em,$$

and

$$\frac{\partial \tilde{w}}{\partial s} = -(w - k - o + m)e' - (1 - e)k' - em'.$$

The FOC of the maximization of social welfare with respect to s is given by

$$\begin{aligned} &- \mathbb{E}\left[(1-e)k' + em' \,\middle|\, \mathcal{B}\right] - \mathbb{E}\left[e'(w - k - o + m) \,\middle|\, \mathcal{B}\right] \\ &+ \frac{\operatorname{Cov}\left[\tilde{u}' + e'p, \tilde{w} \,|\, \mathcal{M}\right] + \mathbb{E}(e\tilde{w} \,|\, \mathcal{M})\, \mathbb{E}(\tilde{u} + e'p \,|\, \mathcal{M})}{MS} = 0, \end{aligned}$$

which gives the result of Proposition 2.

Welfare maximization with respect to p:

We apply the result of Appendix A-3 given in (31) to the function $z = \tilde{w}$, with

$$\frac{\partial W}{\partial p} = 0,$$

which gives the FOC of welfare maximization with respect to p:

$$\mathbb{E}\left[(1-e)\tilde{w} \,|\, \mathcal{M}\right] = 0.$$

A special case with $e = 0$:

If $e = 0$, as in Veiga and Weyl (2016), the FOC of welfare maximization with respect to p is given by $\mathbb{E}\left[\tilde{w} \,|\, \mathcal{M}\right] = 0$.

Appendix C-2: Welfare-maximizing selection technology with the de-risking strategy

If the selection technology is chosen such that there is no demand from non-compliant consumers, from (4), since $w_c(b) = (1 + \Delta)y$, $o_c(b) = y$ and $r_c = 0$, the social planner maximizes

$$\int\limits_{\mathcal{B}_c} ((1 - e_c(s))\Delta y - e_c(s)\lambda t_c - a_c(s)) f(b) db \tag{39}$$

We have $\mathcal{B}_c = [0, b_c]$, with $b_c = (1-e_c(s))(\Delta y - p)$. Therefore, social welfare is decreasing with p. To maximize social welfare with the de-risking strategy, the social planner sets the minimum price so that non-compliant consumers are excluded. Therefore, it chooses the price $\hat{p}(s) = \Delta(B+y-b_\phi(s))$. The welfare-maximizing quality of the selection technology with the de-risking strategy is given by:

$$\frac{dW}{ds} = \frac{\partial W}{\partial p}\frac{d\hat{p}(s)}{ds} + \frac{\partial W}{\partial s}.$$

Since $\hat{p}(s) = \Delta(B + y - b_\phi(s))$ and $\Delta\partial b_\phi/\partial s = \Delta\partial b_{nc}/\partial s = -B\Delta D^{'}_{nc}$, we have

$$d\hat{p}(s)/ds = B\Delta D^{'}_{nc},$$

with $D^{'}_{nc} < 0$. If $e^{'}_c > 0$, if there is an interior solution to the welfare-maximization problem, since $\frac{\partial b_c}{\partial p} = -(1 - e_c)$, $\frac{\partial b_c}{\partial s} = -e^{'}_c(\Delta y - p)$, and $w_c(b) - y = \Delta y$, at $p = \hat{p}(s^w)$, we have

$$(-B\Delta(1 - e_c(s^w))D^{'}_{nc} - e^{'}_c(s^w)\Delta(b_\phi(s^w) - B))\tilde{w}_c f(b_c) - (a^{'}_c + e^{'}_c(\Delta y + \lambda t_c))F(b_c) = 0.$$

The first term represents the effect of an increase in the quality of the selection technology on the participation of compliant buyers in the higher-quality market, given that the latter pay the price that is required to exclude non-compliant consumers. It corresponds to the social sorting effect. The second term is the direct cost effect, the third term is the social exclusion effect.

For the exclusion of non-compliant consumers to be possible with the inclusion of compliant ones, we showed in Appendix B-3 that the quality of the selection technology should satisfy to condition (33).

Naturally, the social planner's problem may not admit an interior solution. This happens for instance if $e^{'}_c > 0$, $D^{'}_{nc}$ small and a_c constant. In this case, social welfare is decreasing with the quality of the selection technology. Therefore, the social planner chooses the minimal quality of the selection technology that is compatible with the participation of the monopoly and the conditions given in (33).

Appendix D: Connection with the literature on algorithms

To connect our results with the literature on algorithms, the loss function of Proposition 5 can be decomposed as a sum of different costs for errors.

- **(1) The social cost of errors for a given price:** The monopoly does not choose the socially optimal performance of the algorithm for a given price p^w of the service, which generates a social cost for errors, either when non-compliant consumers buy the service (The False Negative Rate, FNR) or when compliant consumers are excluded (The False Positive Rate, FPN). The social cost of the errors for a given price is the following

  $$\mathbb{E}\left[(1-e)(w-\lambda r-o)-\lambda et-a\,|\mathcal{B}^w(s^w,p^w)\right]-\mathbb{E}\left[(1-e)(w-\lambda r-o)-\lambda et-a\,|\mathcal{B}^m(s^m,p^w)\right].$$

  From Proposition 4, the social cost of errors for a given price is the sum of the costs of the direct cost effect, the sorting distortion and the exclusion distortion for a given price.

  If we consider that the cost of collecting consumer data a is equal to zero, that the cost t_c of type I errors is also equal to zero and, with a first approximation, that the volume of consumers $\mathcal{B}_i^w(s^w,p^w)$ and $\mathcal{B}_i^m(s^m,p^w)$ are identical and equal to N_i^0 for each category of consumers $i\in\{c,nc\}$, with constant weights w_{nc} and o_{nc}, the sum of these two costs would be equal to:

  $$(e_{nc}(s^m)-e_{nc}(s^w))(w_{nc}-\lambda r_{nc}-o_{nc})N_{nc}^0+(e_c(s^m)-e_c(s^w))(w_c-y)N_c^0,$$

  which is the usual approximation when evaluating the performance of an algorithm with social weights, with respect to a reference point, which is generally assessed with $e_{nc}(s^w)=1$ and $e_c(s^w)=0$. Since $w_c=(1+\Delta)y$, we obtain

  $$-FNR^m(w_{nc}-\lambda r_{nc}-o_{nc})+FPR^m(\Delta y).$$

- **(2) The social cost of errors that is caused by the price distortion:** The monopoly does not choose the socially optimal price of its service for a given quality s^m of the selection technology, which changes the participation of consumers. This additional social cost equals

$$\mathbb{E}\left[(1-e)(w-\lambda r-o)-\lambda et-a\,|\mathcal{B}^m(s^m,p^w)\right]-\mathbb{E}\left[(1-e)(w-\lambda r-o)-\lambda et-a\,|\mathcal{B}^m(s^m,p^m)\right].$$

The social cost of errors for a given price is the sum of the costs of the sorting distortion and the distortion caused by the difference between the private choice of the price of the service and the socially optimal price. Moreover, there is also a cost caused by the variation of the direct cost effect, when the participation of consumers changes. These social costs of errors are usually neglected when evaluating the performance of a classification algorithm.

An example of welfare loss with the de-risking strategy:

Figure 6: an illustration of the welfare loss with the De-risking strategy

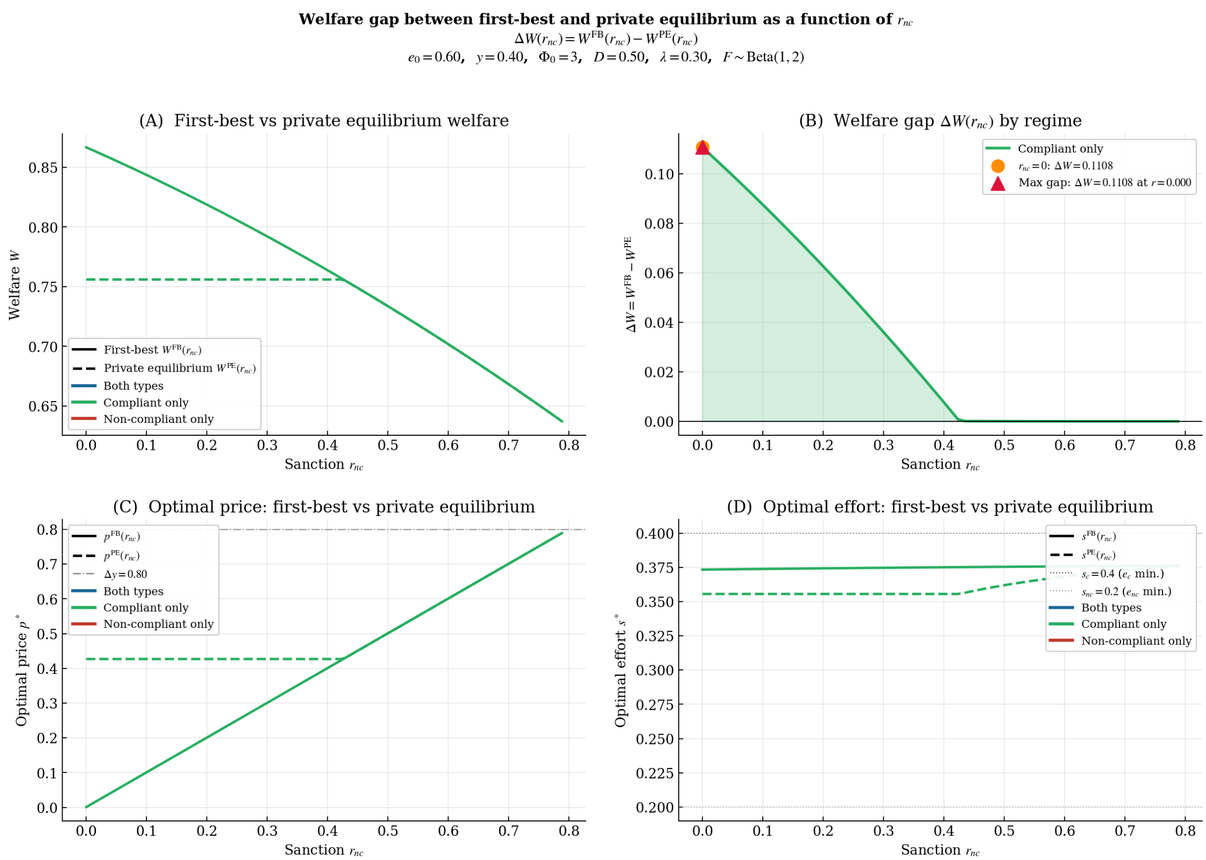


Table 2: Parameters so that both the regulator and the monopoly choose to de-risk:

Component / Parameter	Functional form or definition	Value / notes
Detection functions		
Compliant	$e_c(s)=e_0+a_c(s-s_c)^2-a_cs_c^2$	*U-shaped;* $e_c(0)=e_0$*, minimum at* s_c
Non-compliant	$e_{nc}(s)=e_0+a_{nc}(s-s_{nc})^2-a_{nc}s_{nc}^2$	*U-shaped;* $e_{nc}(0)=e_0$*, minimum at* s_{nc}
Monotonicity	$\frac{d}{ds}[e_{nc}-e_c]=2(a_{nc}-a_c)s-2(a_{nc}s_{nc}-a_cs_c)>0$	*Satisfied for all* $s\geq 0$
Baseline parameter values		
Demand / surplus	$\Delta=2,\ B=1,\ y=0.40$	$\Delta y=0.80$*: compliant participation ceiling*
Detection technology	$e_0=0.60,\ s_c=0.40,\ a_c=0.50,\ s_{nc}=0.20,\ a_{nc}=0.80$	$e_c(1)=0.70,\ \ e_{nc,\min}=0.568$
Baseline intensity	$\Phi_0=3$	$\Phi_0>\Phi_m$*: ib regime structurally blocked*
Detection costs	$c_c=0.05,\ \ c_{nc}=0.075$	
Welfare / damage	$\lambda=0.30,\ \ D=0.50,\ \ d=0.10$	$D^*=(\Delta y+d)/(1+\Delta)=0.30$
Distribution	$\alpha=1,\ \ \beta=2$	$F\sim\text{Beta}(1,2),\ \ \mathbb{E}[b]=1/3$

Appendix E: Proof of Proposition 6

(E-1) Implementation of the de-risking strategy:

The de-risking profit is independent of r_{nc}, whereas the profit with the no-exclusion strategy or the imperfect blocking strategy are increasing with r_{nc}, respectively. Therefore, for any $t_c \geq 0$, there exists a minimum level for the sanction denoted by $r^{ne}(t_c)$ such that the monopoly prefers the de-risking strategy to the no-exclusion strategy. Similarly, there exists a minimum level for the sanction denoted by $r^{ib}(t_c)$ such that the monopoly prefers the de-risking strategy to the imperfect blocking strategy. If the social planner chooses $r_{nc} \geq \underline{r}(t_c) \equiv min\{r^{ne}(t_c), r^{ib}(t_c)\}$, it implements the de-risking strategy. We conclude that for any positive value t_c of the sanction for type I errors, it is possible to implement the de-risking strategy by choosing a minimum value for the sanction for type II errors.

(E-2) Implementation of the imperfect blocking strategy:

If the social planner wishes to implement the imperfect blocking strategy, it should decrease the sanction r_{nc} so that the monopoly prefers the imperfect blocking strategy to the de-risking strategy. However, at the same time, the social planner needs to increase the sanction r_{nc} so that the monopoly prefers the imperfect blocking strategy to the no-exclusion strategy. This sometimes creates a conflict so that the sanction for type II errors is not sufficient to implement the imperfect blocking strategy.

We give the conditions so that the social planner can implement the imperfect blocking strategy with a sanction for each type of error.

- (1) The social planner should choose r_{nc} so that the monopoly prefers the imperfect blocking strategy to the de-risking strategy. From the analysis above (in E-1), this gives $r_{nc} \leq r^{ib}(t_c)$.
- (2) The social planner should choose t_c so that the monopoly prefers the imperfect blocking strategy to the no-exclusion strategy.

Since the no-exclusion strategy is independent of the sanction for type I errors, for any $r_{nc} \geq 0$, there exists a level of sanction for type I errors that we denote by $t^{ne}(r_{nc})$ so that the monopoly makes exactly the same profit for a given r_{nc} with the strategy $j = ne$ or the strategy $j = ib$. Since the monopoly's profit with the imperfect blocking strategy is decreasing with t_c, the monopoly must choose $t_c \leq t^{ne}(r_{nc})$. By the implicit function theorem, it is possible to show that t^{ne} is strictly increasing with r_{nc}. We denote by $(t^{ne})^{-1}$ the inverse function of t^{ne}.

Combining (1) and (2), it is possible to implement the imperfect blocking strategy if and only if there exists $t_c \geq 0$ and $r_{nc} \geq 0$ such that $r_{nc} \leq r^{ib}(t_c)$ and $t_c \leq t^{ne}(r_{nc})$, or else:

$$(t^{ne})^{-1}(t_c) \leq r_{nc} \leq r^{ib}(t_c).$$

A necessary condition for the imperfect blocking strategy to be implementable is that there exists $t_c \geq 0$ such that

$$(t^{ne})^{-1}(t_c) \leq r^{ib}(t_c).$$

Appendix F: Proof of Proposition 7

The social planner maximizes W with respect to r_{nc} and t_c, anticipating that the monopoly will choose the price $p^m(t, r)$ for the product and the quality $s^m(t, r)$ for the selection technology. To simplify, we tackle jointly the maximization with respect to r_{nc} and t_c and denote them indifferently by σ. Solving for the first-order condition of welfare-maximization with respect to σ gives:

$$\sum_{i\in\{c,nc\}}\int_{\mathcal{B}_i}\frac{\partial\tilde{w}}{\partial\sigma}f(b)db+\sum_{i\in\{c,nc\}}(-1)^l(\frac{\partial b_i}{\partial s}s^m_\sigma+\frac{\partial b_i}{\partial p}p^m_\sigma)\tilde{w}=0,$$

with $l=0$ if the consumer is compliant and $l=1$ otherwise. Since

$$\frac{\partial\tilde{w}}{\partial\sigma}=-(e'(w-\lambda(r-t)-o)+a')s^m_\sigma-(1-e)k_\sigma-em_\sigma,$$

this implies that:

$$\frac{\partial W}{\partial\sigma}\frac{1}{q}=M(\mathbb{E}\left[(\tilde{u}'+e'p)\tilde{w}s^m_\sigma\left|\mathcal{M}\right.\right]+\mathbb{E}\left[(1-e)\tilde{w}p^m_\sigma\left|\mathcal{M}\right.\right])$$

$$-q\mathbb{E}\left[(e'(w-\lambda(r-t)-o)+a')s^m_\sigma\left|\mathcal{B}\right.\right]-q\mathbb{E}\left[(1-e)k_\sigma+em_\sigma\left|\mathcal{B}\right.\right].$$

After a division by $q>0$, replacing q/M with MS, we obtain that the per-consumer marginal effect of an increase in the regulatory sanction σ on the additional welfare is:

$$\frac{\partial W}{\partial\sigma}\frac{1}{q}=\frac{1}{MS}(\mathbb{E}\left[(\tilde{u}'+e'p)\tilde{w}s^m_\sigma\left|\mathcal{M}\right.\right]+\mathbb{E}\left[(1-e)\tilde{w}p^m_\sigma\left|\mathcal{M}\right.\right])$$

$$-\mathbb{E}\left[(e'(w-\lambda(r-t)-o)+a')s^m_\sigma\left|\mathcal{B}\right.\right]-\mathbb{E}\left[(1-e)k_\sigma+em_\sigma\left|\mathcal{B}\right.\right],$$

with $k_{t_c}=0$, $k_{r_{nc}}=\lambda$, $m_{t_c}=\lambda$ and $m_{r_{nc}}=0$.

Appendix G: Measures of algorithmic performance

In practice, there exists different performance metrics in order to evaluate algorithms. Here, we consider a binary classification case, as the consumer can be compliant or not to law and regulations. Some of the commonly used metrics are accuracy, sensitivity (also denoted recall), derived from the confusion matrix (here, 2×2).

predicted vs. true	c	nc
c	TP	FP
nc	FN	TN

Table 1: Confusion matrix, with TP= True Positive, TN=True Negative, FP=False Positive, and FN=False Negative the numbers of the predictions with these designations (Rainio et al., 2024).

From these figures in table 1, the following metrics can be calculated:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$
$$Sensitivity = \frac{TP}{TP+FN}$$

In our setting, an improvement in quality s increases the probability of excluding a non-compliant e_{nc} (assumption A.1). In other words, as the quality increases, the number of False Negatives FN should decrease, inducing higher accuracy, and higher recall.

The receiver operating characteristic (ROC) curve can also provide an assessment of the precision of the algorithm. The ROC curve represents the relationship between e_c and e_{nc}, as depicted on Figure 1. To recall, e_c is the probability to exclude a compliant consumer, that is a type I error (false positive rate). $1 - e_{nc}$ is the probability not to exclude a non compliant consumer, that is a type II error (false negative rate). For a given quality, e_{nc} is expressed as a function of e_c, which is the maximum probability of excluding a non compliant consumer given a type I error probability of excluding a compliant one.[52] When the quality s of the algorithm increases, the algorithm is able to obtain a higher e_{nc} given any e_c (assumptions A.1 and A.2). The closer the curve to $(0, 1)$, the more precise is the algorithm. On figure 2, curves in lighter gray refers to a more precise algorithm.

[52] For an analysis of accuracy and legal errors, see Mungan et al. (2023) and Fluet and Mungan (2025).

Figure 7: Type I and type II errors (Mungan et al., 2023)

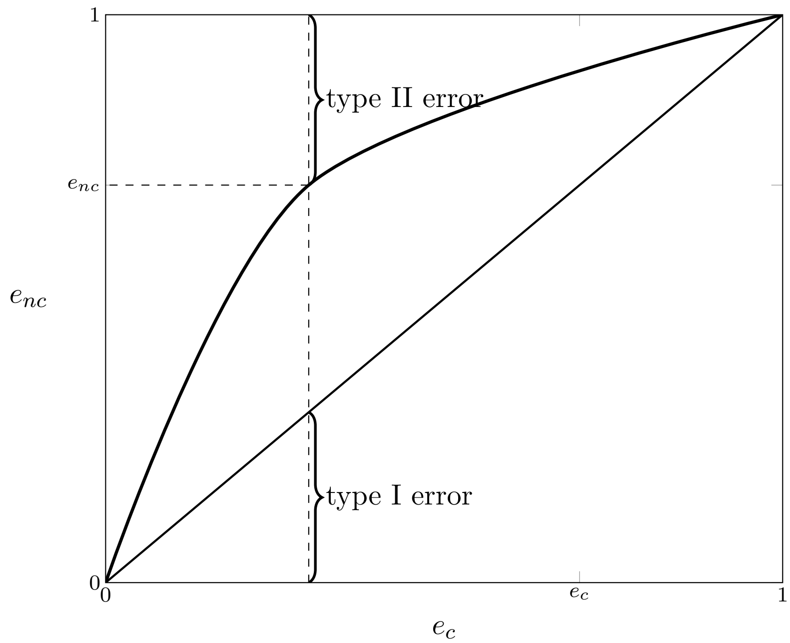


Figure 2: Relationship between quality s, type I and type II errors Mungan et al. (2023)