CRED WORKING PAPER n^o 2025-04

Regulation of Selection Technologies

April, 2025

MARIE OBIDZINSKI* MARIANNE VERDIER†

*Université Paris-Panthéon-Assas, CRED, France.
†Université Paris-Panthéon-Assas, CRED, France.

Regulation of Selection Technologies*

Marie Obidzinski† Marianne Verdier‡

April 9, 2025

Abstract

We analyze how a monopoly chooses the quality of a technology to select its consumers, when it is uncertain that the latter will comply with the legal rules required to buy its product. The firm may decide to exclude a consumer after observing a signal received on her compliance, which accuracy depends on the quality of the technology. The choice of the selection technology also impacts the consumers' incentives to comply with the legal rules. The firm incurs heterogeneous costs of serving compliant and non-compliant consumers, respectively. We explain why the firm's choice of the quality of the technology differs from the social optimum by extending the model of Veiga and Weyl (2016). Then, we use our model to assess whether the implementation of a fraud detection algorithm improves welfare. We analyze the role of several regulatory instruments to improve social welfare: the regulation of the selection technology, the *ex ante* regulation of the misclassification cost through sanctions, the *ex post* imposition of fines on non-compliant consumers.

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

JEL classification: K4, L51, 031.

*We thank Christophe Hurlin, the participants to the 40th symposium on Money, Banking and Finance of the GDRE conference (July, 2024), 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 EA 7321, 75005 Paris, France. e-mail: marie.obidzinski@u-paris2.fr

‡Paris-Panthéon-Assas University, CRED EA 7321, 75005 Paris, France. e-mail: marianne.verdier@u-paris2.fr

1 Introduction

In selection markets, firms often incur higher costs of selling to some consumer types which are costly to observe. To segment the market, they sometimes successfully offer menus of contracts with different prices and qualities. Then, consumers self-select themselves by choosing the contract corresponding to their type (Stiglitz, 1977, Mussa and Rosen, 1978). However, in many cases, firms prefer instead to offer a single contract to their consumers, with the use of non-price features (e.g., downpayment, collateral) to sort their consumers (Rothschild and Stiglitz, 1976, Veiga and Weyl, 2016). A common instrument to sort consumers is the choice of a *selection technology*, which predicts a classification of consumers. Given the result of the prediction, the firm may decide to exclude some consumers from the market.

The choice of a selection technology is particularly widespread in markets where consumers have to comply with legal rules 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 some funding. Consumer compliance is often costly to enforce, because the firm has to invest in a technology to try and identify the consumer's type. A firm may decide to sell to non-compliant consumers, but at a higher cost, either in terms of reputation or regulatory sanctions. When it uses an imperfect selection technology to filter its consumers, the firm may sell at a uniform price, while incurring the costs of making errors, which differ according to the consumer's type. This implies that consumers differ in their imperfectly-contractible profitability to the firm.

When a firm uses a selection technology, it may still partly control who buys its product by changing prices (Akerlof, 1970, Einav and Finkelstein, 2011). Veiga and Weyl (2016) label this effect as *selection by quantity*.[1] However, the firm also relies 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-

[1] Unlike Veiga and Weyl (2016), we consider that product quality is exogenous and the choice of the quality of a selection technology is endogenous. The quality of the selection technology also impacts the consumer's decision to buy the product.

pay for the product for a given price according to their decision to comply with the law. This is because consumers may anticipate that, with some probability, the firm will prevent them from purchasing its product. Therefore, in such markets, firms also rely on *sorting by the quality* (Veiga and Weyl, 2016), through the endogenous choice of the quality of the selection technology.

Our paper offers an economic framework for analyzing the welfare efficiency of a detection algorithm, when the firm is able to select consumers by quantity and by the quality of the screening technology. For this purpose, we analyze how a monopolistic firm chooses the quality of its selection technology and why this choice may not be socially optimal. We offer a measure of the welfare effects of the choice of a detection technology that takes into account the firm's trade-off between selecting by quantity and sorting by quality of the selection technology.

Our model can be applied to fraud detection algorithms, which a firm may use to estimate the probability that a buyer is compliant and then refuse to sell her the product. The firm trades off the profits of expanding its sales to non-compliant buyers and incurring the expected costs of the regulatory sanctions. We consider the example of a bank's anti-money laundering (AML) screening process. The bank offers consumers the convenient option to make electronic payments from their account instead of paying in cash. However, criminals cannot open a bank account because money laundering is forbidden. The bank does not observe consumer types and needs to screen its consumers with a Know-Your-Customer policy. In case the bank does not perform its due-diligence duty and allows a criminal to open an account, it has to incur the expected cost of sanctions, which are decided by the supervisor. This implies that the bank incurs different expected costs in serving heterogeneous types of consumers. A better-quality selection technology may deter criminals from engaging in illegal activities.

We model a market in which consumers receive some benefit of renouncing to comply with the law. In this market, complying with the law is necessary to buy a higher-quality product or service offered by a seller. However, the monopolistic seller does not observe a consumer's type, and needs to rely on a screening technology, which 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. Anticipating 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. After this segmentation, a consumer's type is bi-dimensional, because it is both characterized by its benefit of not-complying with the law and its effective choice of becoming non-compliant, respectively. Non-compliant consumers may sometimes have a higher willingness-to-pay for the firm's services, which generates adverse selection. The firm may select its consumers by increasing its price or the quality of the selection technology.

We determine how the firm chooses the quality of its selection technology and why this choice may differ from the social optimum. We identify another effect that extends the framework of Veiga and Weyl (2016). If the quality of the selection technology increases, this changes the marginal probability that a buyer is excluded from the market. However, the firm only takes into account the impact of the marginal variation of the exclusion probability on its cost, without internalizing the social damage. We use our framework to offer a measure of the welfare effects of fraud detection algorithms, which differs from the existing literature. We introduce a loss function for errors that applies more generally to any selection market in which a firm incurs heterogeneous costs of serving different types of consumers.

We complete our analysis by analyzing the role of sanctions, which the supervisor may use to increase the firm's cost of serving non-compliant consumers.[2] The firm passes through the costs of the sanctions into higher prices. Therefore, the regulator sometimes needs to trade off the costs of sanctioning non-compliant firms against the benefits of including compliant consumers in the market.

The rest of the paper is organized as follows. Section 2 surveys the literature that is related to our study. Section 3 develops the model. Section 4 determines the firm's choice of the quality of the selection technology and compares it to the first-best. Section 5 concludes.

[2]We consider the case of monetary fines. We also consider the case where the regulator may choose to withdraw the monopoly's right to operate.

2 Related literature

Our paper offers a model for analyzing the regulation of a monopoly, which undertakes a selection activity. The selection technology impacts consumer expected utility of buying a higher-quality service with respect to an outside option. Therefore, our framework also has similarities with the model of Spence (1975), which analyzes the regulation of the quality offered by a monopoly. Our paper differs because we assume that product quality is exogenous and instead consider the endogeneous choice of the quality of a selection technology. The monopoly's selection activity impacts consumer ability to buy the service and indirectly generates social damage. In this context, the role of the regulation consists of providing the monopoly with incentives to select compliant buyers, who generate lower social damage.

Our paper belongs to the literature on selection markets (Einav et al., 2010, Einav and Finkelstein, 2011, Mahoney and Weyl, 2017) and extends in particular the model of Veiga and Weyl (2016). In a selection market, firms incur different costs of serving heterogeneous types of consumers and need to attract the right users to be profitable. This idea has also been developed in several models from the literature of platform markets (i.e., Veiga et al. (2017) for a one-sided platform, or Biancini and Verdier (2023) for a two-sided platform). While this literature considers the firm's choice of product quality, we model the endogenous choice of the quality of a selection technology. We identify the same sorting effect as in Veiga and Weyl (2016), which is caused by the marginal effect of a better-quality selection technology on consumer incentives to buy the product.

The issue of consumer selection is related to the broader analysis of the increasing role of private firms as gatekeepers.[3] 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. We contribute to a recent literature on platform regulation through liability or

[3]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.

negligence (Creti and Verdier, 2014, Hua and Spier, 2022, 2023) by focusing on a selection market.[4] We analyze a second-best situation, in which the social planner may impact the firm's selection costs through the choice of a sanction.

In a selection market, a firm incurs heterogeneous costs of selecting different types of consumers. Selection costs can be related to a literature that incorporates data as an input in 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 (see, for instance, Gurkan and de Véricourt, 2022). 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 (Acemoglu et al., 2022), and competition (Jones and Tonetti, 2020). In our paper, consumer data are a private asset that is used ex ante by a monopoly to screen its consumers. The consumer's benefit of not complying with the legal framework could be interpreted as the value of hiding some private data, which affects the marginal cost of the monopoly.[5] A consumer's expected utility of buying the product is all the more reduced since she faces a high probability of being excluded during the screening process. This is similar to other papers of the literature on privacy, such as Markovich and Yehezkel (2021), who model the inconvenience cost for consumers of data collection. In our paper, this cost is group-specific and it endogenously depends on the firm's choice of a selection technology. Other articles in the literature on privacy assume that users have heterogeneous preferences for privacy (see O'Brien and Smith, 2014, Jullien et al., 2020, Acemoglu et al., 2022). 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).[6] A different type of externality arises in our framework between consumers when they decide whether or not to comply with the legal framework. Since consumers pay a uniform price, the probability that a non-compliant consumer wishes to buy the higher-quality

[4]These last two theoretical papers are complemented by a paper analyzing US legislation and case law on platform liability (see Spier and Van Loo, 2025).

[5]Unlike Jones and Tonetti (2020) who model data as a non-rival input generated as a by-product of economic activity, data is a private asset in our model.

[6]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.).

product impacts the compliant consumers' utility.

Artificial Intelligence (AI) is an example of a technology that can be used to select consumers. The role of AI consists of screening a higher volume of data more efficiently than humans and helping them make a decision (see Cowgill et al., 2025,Goh and Lee, 2019). The empirical literature on the use of AI in selection markets focuses on the analysis of a specific algorithm, or compares the efficiency of several methods or criteria to select consumers (see Fraisse and Laporte, 2021, and Hurlin et al., 2024 for credit scoring or Zhang and Trubey, 2019 for AML). These empirical papers do not analyze the interactions between the firm's choice of an algorithm and its strategy in the product market. In particular, the usual measures of algorithmic performance such as accuracy or AUC fail to account that the quality of the screening technology affects consumer demand.

We contribute to a scarce theoretical literature that focuses on the interactions between AI adoption and product pricing. In a close paper, Gans (2022) analyzes AI adoption by a monopoly facing demand uncertainty. In his paper, the firm chooses its price and quantity ahead of demand and may make two different types of errors: unsold inventory or missed sales. Our paper differs because we consider that technology is used to select consumers, with different consequences for the society in terms of errors. Gurkan and de Véricourt (2022) model the firm's incentives to price its product in a two-period model to collect some data on its consumers so as to feed new data back to the algorithm. They show that the firm has an incentive to underprice the product in the first period to collect more data when an increase in the provider's effort has a significant positive impact on accuracy.

An important aspect of AI adoption is related to the interactions between technology and human judgment in decision-making (Agrawal et al., 2018, Daugherty and Wilson, 2018, Mullainathan and Spiess, 2017, Kleinberg et al., 2017). We abstract from studying these issues by considering that the firm always follows the recommendation of the selection technology ex post but that it may ex ante choose the precision of the signal received. We also assume that all agents are perfectly informed about the selection technology.

Our paper is related to the literature on the regulation of AI (Rambachan et al., 2020, Battiston et al., 2024, Acemoglu and Lensman, 2024). Battiston et al. (2024) offer a method to rank different algorithms with heterogeneous costs of type I and type II errors, but do not consider the interactions of AI regulation and product pricing. They apply their approach to the prediction of tax evasion. Our paper does not consider fairness issues, which are also important in AI regulation. Rambachan et al. (2020) determine how a social planner should choose an algorithm when a group of disadvantaged users exhibits different observable characteristics that may imply their exclusion. Acemoglu and Lensman (2024) show in a dynamic setting that firms tend to adopt technology too quickly when it is possible to learn about the potential negative effects of the technology. As in their paper, we also consider that the firm may not choose the socially optimal adoption of the technology and analyze how sanctions and end-user taxes may correct for misuse of the technology. In contrast, we do not consider multiple sectors and focus on the adoption of a selection technology in a given sector with a static game.

3 The model

We build a model to study a monopoly's choice of the quality of a selection technology. The monopoly may use the result of the prediction delivered by the selection technology to refuse to sell to some consumers, when there is a likelihood that some consumer categories are more costly to serve. The firm chooses the quality of the selection technology, which impacts the segmentation of consumer types. For example, a bank may use a screening technology such as an algorithm to separate honest consumers from money launderers.[7] If it does not detect criminals, the bank incurs the expected cost of the regulatory sanctions, which implies that criminals are more costly to serve than honest consumers. A better-quality screening technology provides consumers with incentives to remain honest.

To proceed, we modify the framework of Veiga and Weyl (2016) of a selection market, and use it to discuss the regulation of a selection technology.[8] These authors build a general model to analyze

[7]The AML in financial institutions is built on two pillars: vigilance at the beginning of a business relationship (including the K.Y.C duties) and constant vigilance throughout the business relationship.

[8]As shown in Appendix A-2, we need to adapt the model of Veiga and Weyl (2016) because our framework does not exactly satisfy to their assumptions.

the choice of product quality by a monopoly when consumer types are multi-dimensional.[9] They show that the monopoly distorts the choice of product quality compared to a social planner, when the firm has the opportunity of sorting consumers by both quantity and quality. We differentiate from their work by considering that product quality is exogenous, whereas the choice of the quality of the selection technology is endogenous. Therefore, the monopoly sorts consumers by quantity and by the quality of the selection technology.

Firms

We consider a market in which firms may sell a service to a continuum of risk-neutral consumers, who differ across their benefit b of taking a private action, which impacts a firm's cost of offering its service. Following a large body of literature on crime deterrence (Becker, 1968), we interpret b as the benefit from not complying with the law or regulations.[10] We assume that b is distributed on $(0, B)$ according to the probability density f, with cumulative distribution F, where $B > 0$.

There are two types of firms in this market: a monopoly, which offers a service of exogenous quality $\Delta > 0$ at a price p, and a competitive fringe of firms, which sells a service of quality $\Delta = 0$ at their marginal cost which is equal to zero.[11] Before selling their service, firms may rely on a costly selection technology such as an algorithm, to determine whether a consumer is compliant. If the selection technology predicts that a consumer is not compliant, a firm may exclude her from the market. If a firm does not use any selection technology, all consumers are able to buy its service.[12] The monopoly is able to screen its consumers with a selection technology, which quality S is chosen in an interval $(0, \overline{S})$ of $\mathbb{R}_+$.[13] The competitive fringe does not use any selection technology. If the monopoly excludes a consumer, she is constrained to buy the lower-quality service from the competitive fringe of firms.

[9]The assumption that consumer types are at least bi-dimensional is essential to obtain their results. In our presentation of the model, we will explain why consumer types are bi-dimensional in our setting.

[10]It could also represent the consumer's benefit of hiding some personal information that increases a firm's cost of data collection. See for instance Markovich and Yehezkel (2021).

[11]So far, in the paper, we do not address firms' incentives to sell differentiated product.

[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 Mackowiak, Matejka, and Wiederholt, 2023, for surveys).

[13]We give in Appendix E common empirical measures of algorithmic performance.

Consumers

Compliant consumers have a willingness to pay for a service of quality $\Delta \geq 0$ given by $(1+\Delta)y$, whereas non-compliant consumers have a willingness to pay given by $(1+\Delta)(y+b)$, where $y > 0$ and $b > 0$ is their benefit of being non-compliant. Consumers make two consecutive decisions:

- They decide whether or not to comply with the legal framework, where $i = c$ if a consumer is compliant or $i = nc$ if she is not compliant. After this decision, the consumer's type is bi-dimensional and we denote it by $\theta = (b, i)$.
- They decide whether or not to buy the service of quality $\Delta > 0$ at a price p from the monopoly or a lower-quality version of quality $\Delta = 0$ from the competitive fringe of firms.

When they choose from which firm to buy, we assume that consumers are perfectly informed about the quality S of the selection technology.[14] In case the monopoly detects a non-compliant consumer, it systematically reports it to the regulator and the consumer has to pay a fine $\phi \geq 0$.[15]

The quality of the selection technology determines the likelihood that a consumer will be able to purchase the higher-quality service.[16] Given the quality S of the selection technology, conditional on the consumer being of type θ, the firm excludes her from the market with probability $e(S,\theta)$, and agrees to sell her the higher-quality service with probability $1-e(S,\theta)$, where $0 \leq e(S,\theta) \leq 1$. With probability $e_c \equiv e(S,b,c)$, the firm excludes a compliant consumer (a false positive consumer) and makes an error of type I, whereas with probability $1 - e_{nc} \equiv e(S,b,nc)$, it sells to a non-compliant consumer (a false negative consumer) and makes an error of type II.[17] If $S = 0$, the monopoly does not exclude any consumer, that is, we have $e_c(0,b,c) = e_{nc}(0,b,nc) = 0$, whereas if $S = \overline{S}$, the monopoly excludes a non-compliant consumer with certainty and never excludes a compliant consumer, that is, we have $e_{nc}(0,b,nc) = 1$ and $e_c(0,b,c) = 0$.

[14] This assumption can easily be relaxed by assuming that only a proportion $\eta \in (0,1)$ of consumers is informed about the technology.

[15] 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.

[16] In practice, there exist different metrics to measure algorithms performance, which are exposed in Appendix E.

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

We formalize two assumptions on the selection technology.

(A1) 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 $\partial(e_{nc} - e_c)/\partial S \geq 0$.

We do not make any assumption about the impact of the selection technology on the probability of excluding a compliant consumer. The function e encompasses the case in which a better-quality screening technology has heterogeneous effects on the probabilities of excluding different types of consumers.[18]

To simplify notations, we will use in the mathematical expressions $i = 0$ when the consumer is compliant and $i = 1$ when she is not compliant and the index $i = c$ when the consumer is compliant, or $i = nc$ when she is not compliant. If she buys the higher-quality service from the monopoly, a consumer of type $\theta = (b, i)$ passes the selection process with probability $1 - e_i$, which implies that she obtains the utility $(y + ib)(1 + \Delta)$ and pays the price p. If she fails to pass the selection process with probability e_i, she buys the lower-quality service from the competitive fringe of firms and pays for the fine ϕ if she is non-compliant, which implies that she obtains $y + i(b - \phi)$. Therefore, a consumer of type $\theta = (b, i)$ expects to obtain the utility $u(S, \theta) - (1 - e_i)p$, where

$$u(S, \theta) \equiv (1 - e_i)(y + ib)(1 + \Delta) + e_i(y + i(b - \phi)). \tag{1}$$

If she buys the lower-quality service from the competitive fringe of firms, she does not need to be compliant if there is no selection technology. Therefore, she obtains the utility

$$v(\theta) \equiv b + y. \tag{2}$$

For $i \in \{c, nc\}$, we denote by $\theta_i = (b_i, i)$ the type of the consumer who is indifferent between

[18] A better-quality selection technology often decreases the probability to exclude compliant consumers. However, sometimes investing in the performance of the algorithm can also increase the probability of excluding compliant ones.

purchasing the monopoly's service and the outside option, and it is implicitly defined by

$$\tilde{u}(S,\theta_i) - (1-e_i)p = 0,$$

where $\tilde{u}(S,\theta) \equiv u(S,\theta) - v(\theta)$. From (1) and (2), we have

$$b_i = \frac{ie_i\phi + (1-e_i)(p-\Delta y)}{(1-e_i)\Delta i + i - 1}, \tag{3}$$

with $i = 0$ when the consumer is compliant and $i = 1$ if the consumer is not compliant. If the monopoly does not screen its consumers (i.e., $S = 0$), no consumer complies when purchasing the higher-quality service. If $S > 0$, we denote by b_I the value of the benefit b that leaves the consumer indifferent between being compliant and not being compliant when she consumes the higher-quality service, and it is implicitly defined by $u(S, b_I, nc) - u(S, b_I, c) = 0$, or else:

$$b_I \equiv \frac{(\Delta y - p)(e_{nc} - e_c) + \phi e_{nc}}{1 + \Delta(1 - e_{nc})}.\text{[19]} \tag{4}$$

We denote by $\mathcal{B}_c$ (resp., $\mathcal{B}_{nc}$) the set of compliant consumers (resp., non-compliant) who prefer to buy the firm's service rather than the outside option if they are not excluded, and by $\mathcal{B} \equiv \mathcal{B}_c \cup \mathcal{B}_{nc}$.[20] The set of consumers of type $i \in \{c, nc\}$ who prefer the outside option rather than the firm's service is denoted by $\mathcal{O}_i$. The set of marginal buyers of type $i \in \{c, nc\}$ is defined as $\mathcal{M}_i \equiv \{b : b = b_i\}$, and the set of marginal buyers is $\mathcal{M} \equiv \mathcal{M}_c \cup \mathcal{M}_{nc}$.

Consumer demand for the higher-quality service depends on the price p of the service and the quality S of the selection technology, and it is given in Appendix A-1. Given the quality of the selection technology S and the price p for the higher-quality service, the firm 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

[19] If $b_I \geq B$, all consumers comply when they purchase the higher-quality service, whereas if $b_I \leq 0$, no consumer complies.

[20] The consumer's benefit of being non-compliant b is equivalent to τ in Veiga and Weyl (2016).

the firm's product is

$$Q(S,p) \equiv D_c(S,p) + D_{nc}(S,p) = \sum_{i\in\{c,nc\}} \int_{\mathcal{B}_i} f(b)db.$$

The costs of the quality of the selection technology

If the quality of the selection technology is S, the monopoly incurs a marginal cost $c(S,\theta,r)$ of serving a consumer of type θ, where $r \geq 0$ represents the potential impact of the regulation on the firm's selection costs. For example, the regulator may choose to impose a sanction $r > 0$ on the firm when it sells to non-compliant consumers.

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 p. Since the total quantity of buyers Q is strictly decreasing in p, there exists a differentiable inverse demand function $P(S,q)$ such that $Q(S,P(S,q)) = q$.[21] The firm's profit is given by:

$$\tilde{\pi} \equiv \sum_{i\in\{c,nc\}} \pi^i(S,q,r), \tag{5}$$

where $\pi^i(S,q,r) \equiv \int_{\mathcal{B}_i} (1-e_i)(P(S,q) - c(S,b,i,r))f(b)db$.

The social planner's objective

We define the gross contribution to welfare of an individual of type $\theta = (b,i)$ who consumes the higher-quality service if she is not excluded as $w(\theta)$ (respectively, $o(\theta)$ if she consumes the outside option).[22] We neither include in w nor in o the consumers' private benefit of renouncing to comply

[21]It is equivalent to maximize the firm's profit with respect to (S,p) or (S,q).

[22]In particular, these contributions include the potential social damage generated by the consumption of non-compliant consumers. Our approach allows for different level 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.

with the legal framework.[23] The social planner maximizes total social welfare given by

$$W(S,p,r) = \sum_{i\in\{c,nc\}} \int_{\mathcal{B}_i} (1-e_i)(w(\theta) - c(S,\theta,r) - o(\theta)) f(b) db + \mathbb{E}(o(\theta)). \tag{6}$$

either with respect to S and p, or with respect to r. We assume that for any quality of the selection technology $S \in (0, \overline{S})$ and for any $b \in (0, B)$, compliant consumers always generate a positive gross contribution to social welfare when they consume the higher-quality service instead of the outside option, that is, we have $w(b,0) - c(S,b,0,r) - o(b,0) > 0$.[24]

Additional notations

Similar to Veiga and Weyl (2016), we define the marginal impact of a price increase on the set of marginal buyers of type $i \in \{c, nc\}$ by

$$M_c(S,p) = -\frac{\partial b_c}{\partial p} f(b_c) > 0, \tag{7}$$

and

$$M_{nc}(S,p) = \frac{\partial b_{nc}}{\partial p} f(b_{nc}) > 0, \tag{8}$$

respectively.[25] Then, using the notations of their paper, we have $M(S,p) = -\partial Q/\partial p$, and $M(S,p) = M_c(S,p) + M_{nc}(S,p)$. The marginal consumer surplus is

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

[23] This approach differs from the literature issued from Becker (1968) (see for instance Polinsky and Shavell, 2007, and Garoupa, 1997). Public law enforcement models, which look at the optimal fine and the resources allocated to detection, generally include the benefits of criminals in social well-being. However, this hypothesis is controversial (Dau-Schmidt, 1990, Lewin and Trumbull, 1990). Furthermore, in our model, this benefit could be included in our framework with a weight λ in the social planner's objective, which would not affect the intuitions of our results. We focus on the special case in which $\lambda = 0$.

[24] Otherwise, analyzing why the provision of a higher-quality service may improve welfare is irrelevant

[25] We use opposite signs because b_c is decreasing with the price p, whereas b_{nc} is increasing with the price p (see Eq.(3)).

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

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

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

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

with $i=0$ or equivalently $i=c$ and $i=1$ or $i=nc$, and M given above.

We use Newton's notation to denote partial derivatives with respect to S, namely,

$$z'(S,b,i) = \frac{\partial z(S,b,i)}{\partial S}.$$

Timing of the game

We consider the following game:

t=0 Nature chooses the magnitude of the private non-compliance benefit b obtained by a consumer.

t=2 Either the firm or the regulator chooses the quality of the selection technology S and the price p for the higher-quality service.

t=3 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.

t=4 Selection stage:

(a) Each individual decides whether or not to consume the firm's product.

(b) The firm decides whether or not to exclude the consumer from the market.

(c) If the firm accepts to sell to the consumer, the consumer pays the price p and obtains the additional value of consuming the service. Otherwise, she consumes the outside option at no cost.

t=5 Law enforcement stage:

(a) The regulator audits the firm's selection process and may fine the firm for failing to exclude non-compliant consumers.

(b) The judge punishes the non-compliant consumers if the latter have been detected by the firm.

4 Regulation of the Selection Technology

In this section, we determine the firm's private choice of the quality of its selection technology. Then, we compare it to the first-best, and to a second-best situation, in which the social planner is able to impact the firm's selection cost. We also determine whether the provision of the higher-quality service is socially optimal.

4.1 Profit Maximization

Both the price of the higher-quality service and the quality of the selection technology affect consumer demand for the monopoly's service. Since compliant consumers receive an additional benefit Δy of buying the higher-quality service instead of the outside option, for their demand to be positive, it must be that $p \leq \Delta y$. If the monopoly chooses $p > \Delta y$, it excludes compliant consumers.

We focus on the case in which the monopoly chooses a price such that there may be some positive demand of compliant consumers. In that case, the quality of the selection technology determines whether the monopoly either sells to one or both consumer categories. Since the demand of non-compliant consumers is decreasing with the monopoly's price p and the quality of the selection technology S, for any price $p \leq \Delta y$, there exists a minimal quality of the selection technology $\tilde{S}(p)$ such that non-compliant consumers do not wish to buy the monopoly's service, and it is implicitly defined by

$$e_{nc}(\tilde{S}(p)) = \frac{\Delta(y+B) - p}{\phi + \Delta(y+B) - p}.^{26} \tag{9}$$

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

- **a no-exclusion technology ($j = ne$):** no selection of consumers, that is, we have $S = 0$, which implies that all consumers are non-compliant,
- **a de-risking technology ($j = dr$):** a selection technology $S \in (\tilde{S}(p), \overline{S})$, which implies that there is no demand from non-compliant consumers at any price chosen by the intermediary,
- **an imperfect blocking technology ($j = ib$):** a selection technology $S \in (0, \tilde{S}(p))$, which may generate a positive demand from both compliant and non-compliant consumers.

We denote the monopoly's maximum profit of choosing strategy $j \in \{ne, dr, ib\}$ by π^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 $\pi^m = \max(\pi^{ib}, \pi^{ne}, \pi^{dr})$, where $m \in \{ne, dr, ib\}$ represents the monopoly's profit-maximizing strategy. The choice of the quality of the selection technology depends on how it impacts the probabilities to exclude compliant and non-compliant consumers, respectively, and the marginal cost of serving each type of consumer. In what follows, we focus on the most interesting case in which the monopoly serves both types of consumers.

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. *Suppose that the monopoly serves both types of consumers. The per-consumer marginal effect of an increase in the quality of the selection technology on the monopoly's profit is:*

$$\frac{\partial \tilde{\pi}}{\partial S}\frac{1}{q} = -\mathbb{E}\left[(1-e)c' \,|\mathcal{B}\right] + \frac{\text{Cov}\left[\tilde{u}' + e'p, (1-e)(P-c) \,|\mathcal{M}\right]}{MS} - \mathbb{E}\left[e'(P-c) \,|\mathcal{B}\right] + \mathbb{E}\left[\tilde{u}' + e'p \,|\mathcal{M}\right] \mathbb{E}\left[1-e \,|\mathcal{B}\right].$$

Proof. See Appendix B-1. □

[26] If $S = \tilde{S}(p)$, we have $b_{nc} = B$, and the demand of non-compliant consumers is null.

An increase in the quality of the selection technology has four effects on the firm's profit. The first three effects are similar to Veiga and Weyl (2016), with the nuance that the firm is able to exclude some consumers from the market. The last effect is specific to our setting.

- 1) **Direct cost effect**: when the firm increases the quality of the selection technology, it loses the average increase in the cost of all buyers *who are not excluded*, which results from an increase in provided quality, $\mathbb{E}\left[(1-e)c' \,|\mathcal{B}\right]$. Veiga and Weyl (2016) obtain the same effect with $e = 0$.

- 2) **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 with a high or a low c. Veiga and Weyl (2016) obtain the same effect, because if $e = 0$, we have $\text{Cov}\left[\tilde{u}' + e'p, (1-e)(P-c)\,|\mathcal{M}\right] = -\text{Cov}\left[\tilde{u}', c\,|\mathcal{M}\right]$. 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).

- 3) **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 that a consumer is excluded increases in average, this represents a cost for the firm, $\mathbb{E}\left[e'(P-c)\,|\mathcal{B}\right]$. This term cancels out when the probability to exclude a consumer is independent of the quality of the selection technology.

- 4) **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 additional quality. The average benefit takes into account the average probability that buyers are not excluded. Veiga and Weyl (2016) obtain the same effect with $\mathbb{E}\left[1-e\,|\mathcal{B}\right] = 1$ and $e' = 0$.

4.2 Welfare Maximization

In this section, we determine the welfare-maximizing choice of the quality of the selection technology and we explain why the monopoly's choice is not socially optimal. There are two different distortions in our setting. First, the monopoly may not choose to sell to the socially optimal set of consumer categories. For example, it may choose to sell to both types of consumers, though a social planner prefers the de-risking strategy with a complete exclusion of non-compliant consumers. Second, even if the monopoly chooses to sell to the socially optimal set of consumer categories, it may not choose the welfare-maximizing quality of the selection technology.

The social planner trades off between choosing a pair (S, p) of quality and prices such that the monopoly sells to both types of consumers, a de-risking strategy such that there is no demand from non-compliant consumers, and a no-exclusion strategy such that no consumer category is excluded. We determine the maximum social welfare for $j \in \{ne, dr, ib\}$ and denote it by W^j. The maximum welfare is $W^w = \max(W^{ib}, W^{dr}, W^{ne})$, where $w \in \{ne, dr, ib\}$ represents the social planner's welfare-maximizing strategy.

To better understand how the monopoly distorts the choice of the quality of the selection technology, we start by assuming that the latter sells to the welfare-maximizing set of consumer categories. If both the social planner and the monopoly prefer not to exclude any consumer category, analyzing the quality of the selection technology is irrelevant. We therefore need to study the case in which both the social planner and the monopoly either prefer to serve only compliant consumers, or both consumer categories.[27]

We choose to focus on the most interesting case in which both the social planner and the monopoly prefer the latter to serve two consumer categories.

4.2.1 First-best choice of a price for the service

Proposition 2 gives the welfare-maximizing price of the service when the monopoly serves both types of consumers.

[27]We do not consider the case in which both the social planner and the monopoly prefer to serve non-compliant consumers.

Proposition 2. *If the social planner also prefers that the monopoly sells to both types of consumers, the marginal impact of an increase in the price p on social welfare is given by:*

$$\frac{\partial W}{\partial p} = \frac{\partial b_c}{\partial p}(1 - e_c(S, b_c, c))\ (w - c - o)|_{(b_c,c)} - \frac{\partial b_{nc}}{\partial p}(1 - e_{nc}(S, b_{nc}, c))\ (w - c - o)|_{(b_{nc},nc)}\,.$$

Proof. See Appendix B-2. □

In several situations, non-compliant consumers always generate negative welfare effects when they consume the higher-quality service, that is, for any $b \in (0, B)$ and $S \in (0, \overline{S})$ we have

$$w(b, 1) - c(S, b, 1, r) - o(b, 1) < 0.$$

Then, in this case, the socially optimal price is chosen such that the marginal expected social benefit of selling the higher-quality service to the marginal compliant consumer equals the marginal social expected cost of selling the higher-quality service to the marginal non-compliant consumer.

In some cases, the social benefit of selling the higher-quality service to non-compliant consumers who have an intermediate benefit of being non-compliant may exceed the social damage created by those consumers. Then, social welfare is decreasing with the monopoly's price, the first-best is achieved when the price of the service is null.[28]

In the special case in which e does not depend on b, the first-best price is similar to Veiga and Weyl (2016) with $e = 0$.

Corollary 1. *If there is an interior solution and if e does not depend on b, the first-best price p^w satisfies to*

$$\mathbb{E}\left[(1 - e)^2(w - c - o)\,\middle|\,\mathcal{M}\right] = 0.$$

[28]To determine the first-best, the social planner does not take into account the monopoly's participation constraint.

4.2.2 Welfare-maximizing selection technology with the imperfect blocking strategy

Proposition 3 extends Veiga and Weyl (2016) to our setting, and shows that the social exclusion effect may differ from the private exclusion effect, when the probability to exclude a consumer depends on the quality of the selection technology.

Proposition 3. *Suppose that the social planner chooses a selection technology of quality S such that there is a positive demand from both consumer types. The per-consumer marginal effect of an increase in the quality of the selection technology on the additional welfare is:*

$$\frac{\partial W}{\partial S}\frac{1}{q} = -\mathbb{E}\left[(1-e)c' \,|\mathcal{B}\right] + \frac{\text{Cov}\left[\tilde{u}' + e'p, (1-e)(w-c-o) \,|\mathcal{M}\right]}{MS} - \mathbb{E}\left[e'(w-c-o) \,|\mathcal{B}\right].$$

Proof. See Appendix B-2. □

An increase in the quality of the selection technology has three effects on social welfare. The first two effects are similar to Veiga and Weyl (2016), with the nuance that the firm is able to exclude some consumers from the market. The third effect is specific to our setting. A fourth effect is present in Veiga and Weyl (2016) but absent in our setting, because we assume that product quality is exogenous.

1. **Direct cost effect**: when the firm increases quality, it loses the average increase in the cost of all buyers *who are not excluded*, which results from an increase in provided quality, $\mathbb{E}\left[(1-e)c' \,|\mathcal{B}\right]$. Veiga and Weyl (2016) obtain the same effect with $e = 0$.

2. **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. Veiga and Weyl (2016) obtain the same effect, because if $e = 0$, we have $\text{Cov}\left[\tilde{u}' + e'p, (1-e)(w-c) \,|\mathcal{M}\right] = -\text{Cov}\left[\tilde{u}', c \,|\mathcal{M}\right]$. In our setting, the contribution of marginal consumers to welfare differs from their contribution to profit. The difference between the private sorting effect and the social sorting effect generates the sorting distortion.

3. **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, $\mathbb{E}\left[e'(w-c-o)\,|\mathcal{B}\right]$. This term cancels out when the probability to exclude a consumer is independent of the quality of the selection technology. The difference between the private exclusion effect and the social exclusion effect generates the exclusion distortion.

4. **Social Spence term**: 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. Veiga and Weyl (2016) obtain this effect which is absent in our setting because product quality is exogenous, whereas the firm chooses the quality of the selection technology. This implies that $\mathbb{E}\left[w'\,|\mathcal{B}\right]=0$.

4.2.3 The welfare-maximizing provision of the higher-quality service

Given the quality of the selection technology and the price of the service, a natural question that arises is whether the provision of the higher-quality service by the monopolistic intermediary improves welfare. If the intermediary does not offer the higher-quality service, all consumers are non-compliant and purchase the outside option. In this case, social welfare is equal to

$$\underline{W} \equiv \int_0^B o(b,1)f(b)db.^{[29]} \tag{10}$$

Proposition 4 derives the condition such that the intermediary's provision of the higher-quality service is socially optimal.

Proposition 4. *The social planner prefers that a monopolistic intermediary enters the market and offers the higher-quality service with the welfare-maximizing selection technology S^w and the first-best price p^w if and only if $W(S^w, p^w, r) > \underline{W}$. Otherwise, the social planner prefers that the monopoly does not offer the higher-quality service.*

[29] Recall that $o(b,1)$ represents the gross contribution of non-compliant consumers to social welfare.

Proof. If the monopoly does not offer the higher-quality service, no consumer is compliant and social welfare is equal to $\underline{W}$. If the monopoly offers the higher-quality service, social welfare is given by $W(S^w, p^w, r)$ in (6). Therefore, social welfare is higher if the monopoly offers the higher-quality service if and only if $W(S^w, p^w, r) > \underline{W}$. □

In the rest of the paper, we assume that the provision of the higher-quality service is socially optimal, that is, that:

(A3) $W(S^w, p^w, r) > \underline{W}$.[30]

Otherwise, the social planner should prevent the monopoly from entering the market and offering the higher-quality service. In that case, analyzing the regulation of the quality of the selection technology is useless.

If the provision of the higher-quality service improves welfare, the regulator may decide to influence the monopoly's choice of the quality of the selection technology. When a private intermediary chooses both the price of the higher-quality service p^m and the quality of the selection technology S^m, its decision to sell the higher-quality service with a selection technology may differ from the first-best. First, we explained after Proposition 3 why the intermediary distorts the choice of the quality of the selection technology compared to the first-best. Second, the set of marginal buyers, which depends on the price of the service and the quality of the selection technology (that is, $\mathcal{B}_i^m$), also differs from the first-best (i.e., $\mathcal{B}_i^w$). Third, the intermediary does not perfectly internalize the social benefits and costs of its choices.

In particular, the monopoly may sometimes choose a quality for the selection technology and a price for the service, which imply that social welfare is reduced, that is, whenever

$$W(S^m, p^m, r) < \underline{W} < W(S^w, p^w, r).$$

In that case, a regulatory option consists of deciding to prevent entry by offering licenses for the

[30]For example, this condition is not satisfied if the monopoly's profit-maximizing strategy is the no-exclusion strategy and non-compliant consumers generate high social damage.

production of the higher-quality service.

Otherwise, the monopoly's provision of the higher-quality service with a selection technology improves welfare, but its choices of quality and price differ from the first-best. The regulator may rely on several instruments, such as a quality standard, which will affect consumer willingness-to-pay for the service or a sanction, which will affect the monopoly's marginal cost of serving non-compliant consumers. To measure the best-regulatory option, it is convenient to evaluate the welfare losses with respect to the first-best.

4.2.4 Comparing the welfare effects of selection technologies

Proposition 5 evaluates the effect of adopting a selection technology of quality S when the price of the service is p, with respect to the first-best.

Proposition 5. *With respect to the first-best, the adoption of an imperfect selection technology of quality S and with a price p generates a welfare loss given by:*

$$L(S,p) = \int_{\mathcal{B}_{nc}} (1 - e_{nc}(S))(c(S,b,1,r) + o(\theta) - w(\theta))f(b)db$$

$$+W^w - \mathbb{E}(o(\theta)) - \int_{\mathcal{B}_c} (1 - e_c(S))(w(\theta) - c(S,b,0,r) - o(\theta))f(b)db. \tag{11}$$

Proof. From (6), we have $W^w - W = L(S,p)$. □

The loss function measures how the imperfection of the selection technology impacts social welfare when there is both selection by quantity and sorting by quality of the selection technology. It is the sum of two different errors. The first type of error corresponds to the social losses caused by the consumption of the higher-quality service by non-compliant consumers (the set $\mathcal{B}^w_{nc}$), because the firm erroneously accepts to include them (see the first line of Eq.(11)). This error corresponds to the social cost of the False Positive Rate (FPR) and is also called an error of type I in selection models.

The second type of error is due to the fact that there is a positive probability that some consumers are not compliant, which depends on the monopolist's price and the selection technology (see the second line of Eq.(11)). If no consumer is compliant, that is, if $D_c(S,p) = 0$, the opportunity cost of not selling to compliant consumers is the maximal social benefit of offering the service to compliant consumers, instead of having a group of non-compliant consumers who purchase the outside option, that is, $W^w - \underline{W}$. If there is a set of B_c compliant consumers who buy the higher-quality service, the expected marginal social benefit of including them is $(1 - e_c(S))(w(b,0) - c(S,b,0,r) - o(b,0))$, which reduces the marginal social cost of having non-compliant consumers in the economy. This error is not identical to an error of type II in selection models, that is, the social cost of the False Negative Rate (FNR). Our model differs because the monopoly may use the selection technology to sort consumers by quality and the price to select consumers by quantity, which reduces the social damage.

The loss function of Proposition 5 has some similarities with the literature on algorithms, which analyzes the social planner's regulation of selection decisions. 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. The loss function that we introduce applies generally for algorithms which are used in selection markets, because the proportion of consumers belonging to a given group is endogenous and both depends on a firm's choice of a price for the higher-quality service and the choice of the quality of the selection technology.[31]

This loss function also offers an economic framework for comparing different selection technologies. In particular, different algorithms may generate different probabilities of excluding non-compliant and compliant consumers, respectively, and imply different incentives for the intermediaries to invest in the performance of the algorithm, depending on the classification and misclassification costs (e.g., by training the algorithm on richer data sets, or choosing to train the algorithm more often). Suppose that a social planer needs to choose between implementing an algorithm A_i with

[31] 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.

a quality S_{A_i} and a price p_{A_i} for the higher-quality service, then, the algorithm A_1 is preferable to A_2 if and only if $L(S_{A_1}, p_{A_1}) < L(S_{A_2}, p_{A_2})$.

4.3 Choice of the sanction for misclassification errors

We now focus on the more realistic case in which the social planner is neither able to choose the quality of the selection technology nor the price of the product. However, before the firm makes its choices, the social planner may impact the cost of selecting different consumer types through the choice of the parameter r. In particular, this happens when the social planner is able to impose a sanction $r > 0$ to the firm for misclassification errors.

We denote by $p^m(r)$ and $S^m(r)$ the price and the quality of the selection technology that maximize the firm's profit for a given r, respectively, and by p_r^m and S_r^m the marginal impact of an increase in r on the price of the service and the quality of the selection technology, respectively. As shown in the Appendix, the signs of p_r^m and S_r^m may be either positive or negative.

Proposition 6 gives the per-consumer marginal effect of an increase in the regulatory parameter r.

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

$$\frac{\partial W}{\partial r}\frac{1}{q} = -\mathbb{E}\left[(1-e)(c'S_r^m + \frac{\partial c}{\partial r})\,\middle|\mathcal{B}\right] + \frac{1}{MS}\mathbb{E}\left[(\tilde{u}' + e'p)(1-e)(w-c-o)S_r^m\,\middle|\mathcal{M}\right]$$

$$-\mathbb{E}\left[e'(w-c-o)S_r^m\,\middle|\mathcal{B}\right] + \frac{1}{MS}\mathbb{E}\left[(1-e)^2(w-c-o)p_r^m\,\middle|\mathcal{M}\right].$$

Proof. See Appendix C. □

The social planner may indirectly impact the choice of the quality of the selection technology by choosing a regulatory parameter r that impacts the cost of selecting consumers, such as a sanction. However, the selection cost may be passed through to consumers into higher prices. Therefore, the social planner needs to take into account how the regulation jointly impacts the price of the service and the quality of the selection technology.

An increase in the regulatory parameter has four effects on social welfare. First, there is a direct cost effect: an increase in r changes the marginal cost of serving consumers who are not excluded. This effect is the resultant of the direct impact of r on the marginal cost of selecting consumers and the indirect impact of r on the intermediary's choice of the quality of the selection technology. Second, there is an indirect impact of r on the social sorting effect. Third, there is an indirect impact of r on the social exclusion effect. Fourth, an increase in r impacts the intermediary's choice of a price for the service, which generates a marginal social benefit and a marginal social cost.

5 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 chooses the quality of a selection technology, that is, the inefficient decision to exclude some consumers from the market. Sometimes, the monopoly does not sell to the welfare-maximizing set of consumer categories. In addition, even if it includes all consumer categories, it does not choose the socially optimal quality of the selection technology. We also offered a measure of the welfare effect of the choice of a fraud detection algorithm in a selection market.

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.

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

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.

Dau-Schmidt, K. G. (1990). An economic analysis of the criminal law as a preference-shaping policy. Duke LJ, page 1.

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

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. (2024). Oriented data-generating processes: A categorization of roc curves. exas A&M University School of Law Legal Studies Research Paper (forthcoming).

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. (2022). Ai adoption in a competitive market. National Bureau of Economic Research Working Paper.

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, 2019(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. (2022). Holding platforms liable. HKUST Bus. Sch. Rsch. Paper no. 2021-048.

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

Hurlin, C., Pérignon, C., and Saurin, S. (2024). The fairness of credit scoring models. Management Science. *Forthcoming.*

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

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.

Lewin, J. L. and Trumbull, W. N. (1990). The social value of crime? International Review of Law and Economics, 10(3):271–284.

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.

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., T. J. K. R. (2024). Evaluation metrics and statistical tests for machine learning. Technical report, Nature. Sci Rep 14.

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.

Rothschild, M. and Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. The Quarterly Journal of Economics, 90(4):629–649.

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. *Forthcoming.*

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

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

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.

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

In this preliminary Appendix, we determine consumer demand for the service according to the price charged by the monopoly. In order to achieve it, we need to determine different thresholds on the benefit to be non-compliant b. Let us denote b_c the type of agent indifferent between being a compliant consumer of the service and being a non-compliant consumer of the outside option, b_{nc} the type of the agent being a non-compliant consumer of the service and a non-compliant consumer of the outside option.

From (1), a compliant consumer who buys the service from the firm obtains utility $u(S, b, 0) \equiv u_c$, with:

$$u_c = (1 - e_c)(y(1 + \Delta) - p) + e_c y$$

The firm agrees to sell the service with probability $1 - e_c$. The agent consumes instead the outside option with probability e_c. A non-compliant consumer buys the outside option at no cost and obtains utility $v(\theta) = y + b$. The indifferent compliant consumer between buying the service and becoming instead a non-compliant consumer who buys the outside option ($u_c = v(\theta)$) is given by

$$b_c \equiv (\Delta y - p)(1 - e_c).$$

If the firm does not exclude any consumer from the market, all consumers are non-compliant and obtain $v(\theta) = b + y$ with certainty. If they buy the service from the firm, they obtain an additional utility $(b + y)\Delta$ but pay the price p. A consumer prefers to buy the service if and only if

$$b \geq b_{ne} \equiv (p/\Delta) - y$$

and prefers the outside option otherwise.

From (1), if the firm excludes some consumers, a non-compliant consumer who buys the service from the firm obtains the expected utility $u(S, b, 1) \equiv u_{nc}$, with

$$u_{nc} = ((b + y)(1 + \Delta) - p)(1 - e_{nc}) + e_{nc}(b + y - \phi).$$

Indeed, the firm agrees to sell with probability $1 - e_{nc}$, and otherwise, with probability e_{nc}, the firm detects the non-compliant consumer, who is constrained to buy the outside option and incur the cost of the fine ϕ. If a non-compliant consumer renounces buying the higher-quality service, he obtains $u_o = y + b$. Therefore, a non-compliant consumer prefers to buy the service ($u_{nc} \geq u_o$) if and only if

$$b \geq b_{nc} \equiv \frac{p}{\Delta} + b_\phi - y,$$

where

$$b_\phi = \frac{\phi e_{nc}}{\Delta(1 - e_{nc})}$$

represents the additional benefit that must be given to the non-compliant consumer, so that the latter adopts the same consumption behavior as if there were no risk of being sanctioned. We have:

$$b_{nc} = b_{ne} + b_\phi.$$

A consumer who buys the service prefers to be non-compliant than compliant ($u_{nc} \geq u_c$) if and only if

$$b \geq b_I \equiv \frac{(\Delta y - p)\hat{e}}{1 + \Delta(1 - e_{nc})} + \frac{\Delta(1 - e_{nc})b_\phi}{1 + \Delta(1 - e_{nc})},$$

where $\hat{e} \equiv e_{nc} - e_c > 0$. The letter I stands for incentives, because the firm's detection technology provides the consumer with incentives to become compliant when the latter obtain sufficiently low benefits of being non compliant. Note that we have $b_I \leq b_c$ if and only if

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

Moreover, 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 $p \leq \underline{p}$.

To summarize, 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 $\mathcal{B}_{nc} = (b_{nc}, B)$.
- If $\underline{p} \leq p \leq \Delta y$, with

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

we have $b_c \leq b_I \leq b_{nc}$. The market is not covered, with $\mathcal{B}_c = (0, b_c)$ and $\mathcal{B}_{nc} = (b_{nc}, B)$.

- If $p < \underline{p}$, we have $b_{nc} \leq b_I \leq b_c$, and the market is covered, with $\mathcal{B}_c = (0, b_I)$ and $\mathcal{B}_{nc} = (b_I, B)$.

Appendix A-2: extending the model of Veiga and Weyl (2016) to our setting

We adapt the definitions given by Veiga and Weyl (2016) to our setting by considering the case in which the market is not covered. 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. First, one of the dimensions of the consumer's type (i.e., i) is a discrete variable. Second, 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, b_c is decreasing with the price p, whereas b_{nc} is increasing with the price p. Third, 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.

Our setting will imply exactly the same selection and sorting effects as in their paper, if we adapt some of their definitions to our framework.

Similar to their paper, from the definition of the marginal buyer of type i given by

$$\tilde{u}(S, b_i, i) - (1 - e_i)p = 0,$$

from the implicit function theorem, we have

$$\frac{\partial b_i}{\partial S} = -\frac{(\tilde{u}'(S, b_i, i) + e'(S, b_i, i)p)}{\frac{\partial \tilde{u}(S, b_i, i)}{\partial b} + \frac{\partial e_i}{\partial b} p}, \tag{12}$$

and

$$\frac{\partial b_i}{\partial p} = \frac{1 - e_i}{\frac{\partial \tilde{u}(S, b_i, i)}{\partial b}}. \tag{13}$$

In our particular setting, we have:

$$\frac{\partial \tilde{u}(S, b_i, i)}{\partial b} = (i\Delta(1 - e_i) + i - 1), \tag{14}$$

and

$$\tilde{u}'(S, b_i, i) = -(\Delta(ib + y) + i\phi)e'(S, b_i, i).$$

Therefore, we have

$$\frac{\partial b_c}{\partial p} = -(1 - e_c), \tag{15}$$

and

$$\frac{\partial b_{nc}}{\partial p} = \frac{1}{\Delta}. \tag{16}$$

Since

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

and from (12), 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, applying the implicit function theorem to the equation $Q(S, P(S, q)) = q$ that defines the inverse demand function yields

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

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:

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

Following Veiga and Weyl (2016), if S is uni-dimensional, the FOC with respect to S is given by:

$$q\mathbb{E}\left[\frac{\partial z}{\partial S}\middle|\mathcal{B}\right] + M\text{Cov}\left[\tilde{u}' + e'p, z\,|\mathcal{M}\right] = 0.$$

Profit maximization considers $z(S,q) = (1 - e(S,\theta))(P(S,q) - c(S,\theta,r))$, whereas welfare maximization considers $z(S,q,r) = (1 - e(S,\theta))(w(\theta) - c(S,\theta,r) - o(\theta))$. Compared to Veiga and Weyl (2016), the function z is multiplied by $(1 - e(S,\theta))$, which represents the probability that the firm does not exclude the buyers of type θ from the market. Veiga and Weyl (2016) consider the special case in which the function e is null, that is, $e = 0$, with $e' = 0$. Moreover, in their setting, the firm chooses product quality, which implies that w may depend on the quality of the service, whereas we consider the quality of the selection technology.

Appendix B-1: Profit Maximization

Profit maximization with respect to S:

When the firm maximizes its profit with respect to S, we have:

$$\frac{\partial z}{\partial S} = -(P(S,q) - c(S,\theta,r))e' + (1 - e(S,\theta))(\frac{\partial P}{\partial S} - c').$$

Since

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

the FOC of the firm's profit-maximization with respect to S is given by:

$$-q\mathbb{E}\left[(1-e)c'\,|\mathcal{B}\right] + q\mathbb{E}\left[\tilde{u}' + e'p\,|\mathcal{M}\right]\mathbb{E}\left[1 - e\,|\mathcal{B}\right] + M\text{Cov}\left[\tilde{u}' + e'p, (1-e)(P-c)\,|\mathcal{M}\right] - q\mathbb{E}\left[e'(P-c)\,|\mathcal{B}\right] = 0,$$

with $q = D_c + D_{nc}$. Replacing for $MS = q/M$ the marginal consumer surplus, after simplification by $q > 0$, the FOC of profit-maximization is given by:

$$-\mathbb{E}\left[(1-e)c'\,|\mathcal{B}\right] + \mathbb{E}\left[\tilde{u}' + e'p\,|\mathcal{M}\right]\mathbb{E}\left[1 - e\,|\mathcal{B}\right] + \frac{\text{Cov}\left[\tilde{u}' + e'p, (1-e)(P-c)\,|\mathcal{M}\right]}{MS} - \mathbb{E}\left[e'(P-c)\,|\mathcal{B}\right] = 0.$$

Profit maximization with respect to p:

From the definitions of M_c and M_{nc} given in (7) and (8), respectively, solving for the first-order condition of the firm's profit maximization with respect to p gives:

$$-\sum_{i\in\{c,nc\}}(1-e_i)(M_i(p-c(S,b_i,i,r))+D_i(S,p))=0.$$

Using the definition of M_c and M_{nc} given in (7) and (8), this implies that:

$$-(1-e_c)\frac{\partial b_c}{\partial p}(p-c)f(b_c)-\frac{\partial b_{nc}}{\partial p}(1-e_{nc})(p-c)f(b_{nc})-(1-e_c)D_c-(1-e_{nc})D_{nc}=0.$$

Replacing for $\frac{\partial b_c}{\partial p}$ and $\frac{\partial b_{nc}}{\partial p}$ given in (15) and (16), respectively, we find that:

$$(1-e_c)^2(p-c)f(b_c)-\frac{(1-e_{nc})^2}{\Delta(1-e_{nc})}(p-c)f(b_{nc})-(1-e_c)D_c-(1-e_{nc})D_{nc}=0.$$

Replacing for $\frac{\partial \tilde{u}}{\partial b}$ given in (14), if e is independent of b, we find that

$$M\mathbb{E}\left[(1-e)^2(p-c)\,\middle|\mathcal{M}\right]-\mathbb{E}\left[(1-e)\,|\mathcal{B}\right]=0.$$

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

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

Appendix B-2: Welfare Maximization

Welfare maximization with respect to S:

When the firm maximizes social welfare with respect to S, we have:

$$\frac{\partial z}{\partial S}=-(w(S,\theta)-c(S,\theta,r)-o(\theta))e'+(1-e(S,\theta))(w'-c').$$

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

$$-q\mathbb{E}\left[(1-e)c' \,\middle|\mathcal{B}\right] + M\mathrm{Cov}\left[\tilde{u}' + e'p, (1-e)(w-c-o) \,\middle|\mathcal{M}\right] - q\mathbb{E}\left[e'(w-c-o) \,\middle|\mathcal{B}\right] = 0.$$

A division of this expression by $q > 0$ and the replacement of q/M with MS gives the result of Proposition 2.

Welfare maximization with respect to p:

Taking the derivative of W with respect to p gives

$$\frac{\partial W}{\partial p} = \frac{\partial b_c}{\partial p}(1 - e_c(S, b_c, c))\,(w-c-o)|_{(b_c,c)} - \frac{\partial b_{nc}}{\partial p}(1 - e_{nc}(S, b_{nc}, c))\,(w-c-o)|_{(b_{nc},nc)}\,.$$

Replacing for $\frac{\partial b_c}{\partial p}$ and $\frac{\partial b_{nc}}{\partial p}$ given in (15) and (16), respectively, if e is independent of b, we find that if there is an interior solution, the FOC of social welfare maximization with respect to p is given by:

$$-\mathbb{E}\left[(1-e)^2(w-c-o)\,\middle|\mathcal{M}\right] = 0.$$

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

Since $\frac{\partial b_c}{\partial p} = -(1-e_c)$ and $\frac{\partial b_{nc}}{\partial p} = \frac{1}{\Delta}$, if $(w-c-o)|_{(b_{nc},nc)} \geq 0$, we have

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

which implies that social welfare is decreasing with the price of the service. Then, the welfare-maximizing price of the service is null, if the social planner does not take into account the monopoly's participation constraint. Therefore, the condition $(w-c-o)|_{(b_{nc},nc)} < 0$ is necessary to obtain an interior solution. If $(w-c-o)|_{(b_{nc},nc)} < 0$ and there is an interior solution, the first-order condition of social welfare maximization is given by

$$\frac{\partial b_c}{\partial p}(1 - e_c(S, b_c, c))\,(w-c-o)|_{(b_c,c)} = \frac{\partial b_{nc}}{\partial p}(1 - e_{nc}(S, b_{nc}, c))\,(w-c-o)|_{(b_{nc},nc)}\,,$$

which means that the marginal expected social benefit of selling the higher-quality service to the marginal additional compliant consumer equals the marginal expected social cost of selling it to the marginal additional non-compliant consumer.

In case the welfare-maximizing price is null, if there is no interior solution, since we assumed that a better-quality selection technology provides consumers with incentives to be compliant, the first-best quality of the selection technology should be maximal. In this case, the entire population of consumers purchases the higher-quality service and is compliant. Therefore, social welfare is given by:

$$\overline{W} = \int_0^B (w(b,0) - c(\overline{S}, b, 0, r)) f(b) db. \tag{17}$$

Appendix C: Second Best:

The social planner maximizes W with respect to r, anticipating that the monopoly will choose the price $p^m(r)$ for the product and the quality $S^m(r)$ for the selection technology. We denote by

$$S_r^m \equiv (S^m)'(r)$$

and

$$p_r^m \equiv (p^m)'(r).$$

Solving for the first-order condition of welfare-maximization with respect to r gives:

$$\frac{\partial W}{\partial r} = \sum_{i \in \{c,nc\}} \int_{\mathcal{B}_i} (-(e_i)'(w(\theta) - c(S^*, \theta, r) - o(\theta)) - c'(1-e_i))S_r^m - (1-e_i)\frac{\partial c}{\partial r}) f(b) db$$

$$+\frac{\partial b_c}{\partial r} (1-e_c)(w-c-o)|_{b_c} - \frac{\partial b_{nc}}{\partial r} (1-e_{nc})(w-c-o)|_{b_{nc}} = 0$$

Since

$$\frac{\partial b_i}{\partial r} = \frac{\partial b_i}{\partial p} p_r^m + \frac{\partial b_i}{\partial S} S_r^m,$$

we have

$$\frac{\partial b_i}{\partial r} = \frac{1-e_i}{\frac{\partial \tilde{u}(S,b_i,i)}{\partial b}} p_r^m - \frac{(\tilde{u}'(S,b_i,i)+e'(S,b_i,i)p)}{\frac{\partial \tilde{u}(S,b_i,i)}{\partial b}} S_r^m.$$

This implies that

$$\frac{\partial W}{\partial r} = M\mathbb{E}\left[(\tilde{u}'+e'p)(1-e)(w-c-o)S_r^m \,\middle|\mathcal{M}\right] + M\mathbb{E}\left[(1-e)^2(w-c-o)p_r^m \,\middle|\mathcal{M}\right]$$

$$-q\mathbb{E}\left[e'(w-c-o)S_r^m \,\middle|\mathcal{B}\right] - q\mathbb{E}\left[(1-e)(c'S_r^m + \frac{\partial c}{\partial r}) \,\middle|\mathcal{B}\right].$$

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 parameter r on the additional welfare is:

$$\frac{\partial W}{\partial r}\frac{1}{q} = \frac{1}{MS}\mathbb{E}\left[(\tilde{u}'+e'p)(1-e)(w-c-o)S_r^m \,\middle|\mathcal{M}\right] + \frac{1}{MS}\mathbb{E}\left[(1-e)^2(w-c-o)p_r^m \,\middle|\mathcal{M}\right].$$

$$-\mathbb{E}\left[e'(w-c-o)S_r^m \,\middle|\mathcal{B}\right] - \mathbb{E}\left[(1-e)(c'S_r^m + \frac{\partial c}{\partial r}) \,\middle|\mathcal{B}\right].$$

Appendix D: 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, social welfare equals

$$W \equiv \int_{\mathcal{B}_c} (1-e_c)(w(\theta)-c(S,\theta,r))f(b)db + \int_{\mathcal{B}_c} e_c o(\theta) f(b)db + \sum_{i\in\{c,nc\}} \int_{\mathcal{O}_i} o(\theta)f(b)db. \qquad (18)$$

To maximize social welfare with the de-risking strategy, since social welfare is decreasing with p, the social planner sets the price of the service at the monopoly's expected cost of serving compliant consumers. The quality of the selection technology that maximizes social welfare with a de-risking strategy depends on how S impacts the probability to exclude compliant consumers and the monopoly's marginal cost of serving compliant consumers. For example, if $c(S,b,0)$ does not depend on S and if e_c decreases with S, the social planner chooses $S^{dr} = \overline{S}$.

Appendix E: 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, 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-1 error probability of excluding a compliant one.[32] 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

[32]For an analysis of accuracy and legal errors, see Mungan et al. (2023) and Fluet and Mungan (2024).

more precise is the algorithm. On figure 2, curves in lighter gray refers to a more precise algorithm.

Figure 1: Type I and type II errors (Mungan et al., 2023)

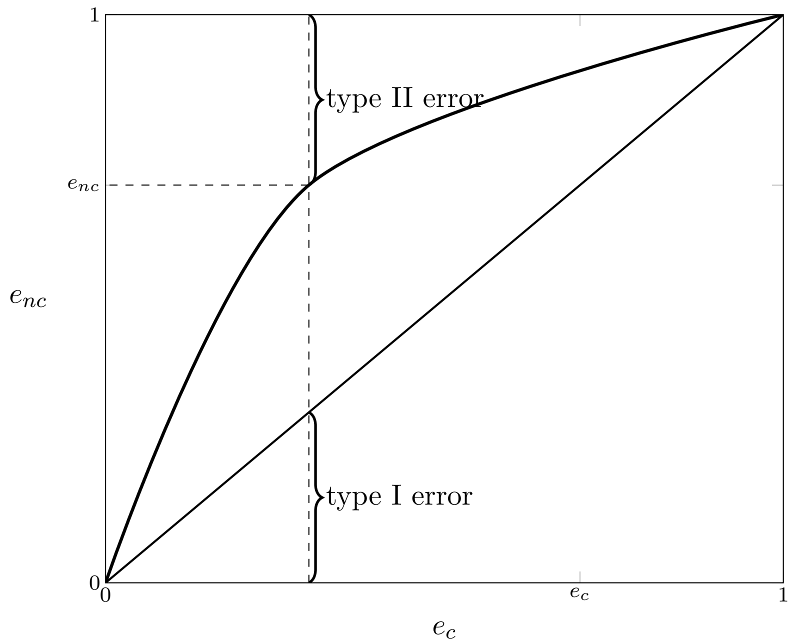

Figure 2: Relationship between quality S, type I and type II errors Mungan et al. (2023)