REGULATING PRIVACY POLICIES ON DIGITAL PLATFORMS

By

Michele Bisceglia, Alessandro Bonatti and Fiona Scott Morton

November 2025

COWLES FOUNDATION DISCUSSION PAPER NO. 2474

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Regulating Privacy Policies on Digital Platforms*

Michele Bisceglia[†] Alessandro Bonatti[‡] Fiona Scott Morton[§]

November 14, 2025

Abstract

We study how privacy regulation affects menu pricing by a monopolist platform that collects and monetizes personal data. Consumers differ in privacy valuation and sophistication: naive users ignore privacy losses, while sophisticated users internalize them. The platform designs prices and data collection options to screen users. Without regulation, privacy allocations are distorted and naive users are exploited. Regulation through privacy-protecting defaults can create a market for information by inducing payments for data; hard caps on data collection protect naive users but may restrict efficient data trade.

JEL Classification: D18, D82, D83, L12, L51.

Keywords: Data, Defaults, Digital Platforms, Menu Pricing, Privacy, Regulation.

*Bonatti acknowledges financial support from NSF grant SES 2519401. We thank Paul Heidhues, Giorgio Monti, and Ellen Muir for helpful comments, and Nick Jacobson for excellent research assistance.

[†]Center for Algorithms, Data, and Market Design at Yale (CADMY). E-mail: michele.bisceglia@yale.edu.

[‡]MIT Sloan School of Management. E-mail: bonatti@mit.edu.

[§]Yale School of Management. E-mail: fiona.scottmorton@yale.edu.

1 Introduction

Large digital platforms such as Amazon, Apple, Google, Meta, and TikTok collect and monetize vast amounts of personal data from billions of users every day. These users differ in their valuations of privacy and in their ability to understand the choices available to them.[1] Many platforms have adopted a "pay-or-consent" model that screens consumers based on their (apparent) privacy preferences. Through this menu-pricing approach, privacy-sensitive users choose to pay a monetary fee for a more private, ad-free experience, while privacy-insensitive users continue using a free, ad-supported service that requires extensive data collection for targeted advertising.[2] These arrangements reveal how platforms explicitly monetize privacy trade-offs and segment their user base through pricing and privacy design.

These developments raise several fundamental economic questions about the welfare consequences of trading personal data on digital platforms. First, is the exchange of user data for digital services efficient from a social perspective? Second, how does this efficiency depend on the platform's ability to monetize user data? Third, how do outcomes depend on the degree to which users understand and internalize the privacy implications of data collection? Finally, given these determinants, which regulatory interventions are most effective in improving welfare?

In this paper, we develop a framework to evaluate the regulation of trade between platforms that collect and monetize personal data and the users who provide that data. We consider a setting in which a platform and its users negotiate data access in exchange for service and use this structure to evaluate how different forms of regulation shape market outcomes and affect consumer and platform surplus. We identify two broad roles for privacy regulation. First, it can protect unsophisticated users from exploitative data extraction by limiting the amount of information traded. Second, it can stimulate efficient markets for data, inducing platforms to share some of the surplus from information with their users.

In our model, a monopolist digital platform designs a menu of service options that vary in price and in the amount of data collected. Consumers differ in their valuations of privacy and in their sophistication about data practices. A fraction of consumers is naive: they do not fully understand that their personal data are being collected and monetized and behave as if all service options provide the same level of privacy. The remaining consumers are sophisticated: they are aware of the privacy terms and take them into account when making decisions. The platform cannot directly observe users' privacy valuations or levels of sophistication and must therefore design a menu of price–privacy contracts to screen them. This setup parallels a second-degree price discrimination problem but with a key twist: the inverse of privacy, personal data collection, reduces service quality for users who dislike data sharing and, if data collection is sufficiently extensive, overall

[1] Significant heterogeneity in privacy preferences is documented in the seminal work of Acquisti et al. (2013, 2015) and more recently in Collis et al. (2021) and Chen et al. (2025).

[2] For example, Meta (Facebook and Instagram) recently introduced a paid ad-free subscription in the European Union alongside its usual free, ad-supported service; similarly, TikTok has tested an ad-free paid tier, YouTube offers YouTube Premium (a paid, ad-free experience), X (formerly Twitter) sells various premium subscriptions with reduced advertising, and Snapchat provides a paid Snapchat+ service.

service quality falls so low that the platform must subsidize consumers to induce participation.

Under complete information, the platform observes each user's valuation for privacy and chooses the prices and data collection level to maximize surplus. The first-best menu is strictly monotone: users who value privacy less share more data. Prices adjust accordingly—low-valuation users receive transfers from the platform in exchange for their data, and high-valuation users pay for additional privacy (Proposition 1).

If privacy preferences are private information, the platform-optimal (second-best) menu exhibits countervailing distortions across types: privacy-sensitive consumers receive too little privacy; a positive-measure bunching region arises at a zero-price, intermediate-privacy option; and consumers who care least about privacy receive too much privacy and negative-price options. Therefore, a data market emerges in which some users sell information and others buy privacy (Proposition 2).

Introducing naive consumers—who ignore privacy losses and choose the cheapest option—fundamentally changes this structure. The platform steers these consumers to a version of the service that extracts the maximum amount of data, while reducing or eliminating negative-price offers to reduce payments to naive consumers. Naive consumers are willing to share all their data for free and therefore compensating them is unnecessary. Once the share of naive users is large enough, negative-price offers disappear entirely (Proposition 3). The *laissez-faire* allocation then features even more over-provision of privacy for low-valuation users, with naive users exploited rather than compensated.

We next consider two regulatory instruments motivated by European data protection law: a hard cap that limits the amount of personal data that the platform can collect; and a requirement that the platform offer a free, privacy-protecting default version of the service. A hard cap prevents extreme extraction and protects naive consumers, but when it binds for sophisticated types it suppresses efficient trade and may reduce welfare (Proposition 4). A privacy-enhancing default regulation that mandates a free, more private option shifts the menu toward greater privacy, strictly raises consumer welfare, and, for intermediate fractions of naive users, induces the platform to introduce negative-price options—creating a data market and paying naive users for their data (Proposition 5). When naive users are sufficiently numerous, however, negative-price options are not offered even under a default.

We then characterize the welfare-optimal combination of the two instruments (Proposition 6). When the share of naive consumers is small, the regulator sets a relatively strict default but a looser cap, enabling efficient data trade while constraining exploitation. When this share exceeds a critical threshold, the optimal policy equates default and cap and forgoes data trade.

Finally, we extend the analysis in two directions. Allowing service quality to be endogenous shows that privacy-enhancing regulation lowers the platform's incentive to invest in quality, with the welfare-optimal default internalizing this quality response (Propositions 7–8). Requiring the user's informed consent to move away from the default either turns naive users into sophisticated users, or leads the platform to keep them at the exploitative free default. In the latter case, default regulation with an informed-consent requirement prevents data exploitation (Propositions 9–10).

1.1 Regulation of Personal Data

The motivation for our research questions is the relatively recent privacy regulation in the European Union. Over the last decade, regulators throughout the world have come to appreciate the value of consumers' personal data and the risk of harm to consumers when third parties can use their data without restraint. Platforms tailor advertising to each user according to what that user has been sharing or viewing, making these ads more valuable. Platforms also commonly use customers' personal data to train AI models. These valuable activities require collection of detailed data on user behavior online (e.g., searches, clicks, products placed into cart) and offline (e.g., location, stores visited). Because this type of digital economic activity is mostly new, in many jurisdictions there are few rules regulating how invasive it can be. For example, in the United States a digital business can collect and profit from user data with a simple "consent" button that the user clicks without reading or understanding the pages of legal text that lie behind it.

In addition to straightforward concerns about the efficient and equitable distribution of surplus from safe digital advertising, other regulatory concerns include harm to user mental health from targeted content, harm to privacy from cross-use of personal data by platforms (e.g., while sharing a screen at work a colleague sees that the user is targeted with an ad for cancer medication), and exploitative advertising (e.g., fraudulent financial products advertised to a gambler).

These concerns have led regulators to create new rules concerning the collection, storage, analysis, and use (legally referred to as "processing") of personal data. The EU has adopted perhaps the most well-known data regulation, the General Data Protection Regulation, which took effect in 2018.[3] This law requires that anyone processing personal data must do so under one of the six legally permitted bases in Article 6, one of which is obtaining the user's consent. GDPR also defines "sensitive data" as data relating to a person's health, sexuality, religion, and the like (Article 9(1)). However, because user consent allows such data processing and because regulators have struggled to regulate the choice architecture of consent to make it substantive, the consent route to processing personal data has become a giant loophole. The European Parliament adopted the Digital Services Act, which came fully into force in 2024.[4] Article 26(3) of the DSA prohibits digital platforms from showing ads to users based on profiling using sensitive data as defined by the GDPR *regardless of user consent*, which motivates our focus on a "hard cap" on data sharing.

A second relevant EU law is the Digital Markets Act. This landmark piece of legislation provides for sectoral regulation of "gatekeepers," namely large digital platforms.[5] Article 5(2) of the DMA prohibits a variety of gatekeeper behaviors, including combining or cross-using personal data

[3] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 4 May 2016, 1-88. Retrieved from `https://eur-lex.europa.eu/eli/reg/2016/679/oj`.

[4] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act), 2022 O.J. (L 277) 1.

[5] Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector (Digital Markets Act), OJ L 265, 12 October 2022. Current version retrieved from `http://data.europa.eu/eli/reg/2022/1925`.

between a gatekeeper's different services without user consent. Because the advertising engines of Google and Meta have been designated as a (different) core platform service, effectively *any* digital advertising by these gatekeepers requires consent from the user. However, the ease of obtaining user consent depends on the specifics of choice architecture and we know from existing work on consumer choice that consumers are very sensitive to the specifics of framing, timing, design etc. (Mertens et al., 2022; Farronato et al., 2025). In principle, compliant choice architecture would ensure users have meaningful informed choice. For example, the GDPR Recital 32 describes user choice with these words: "Consent should be given by a clear affirmative act establishing a freely given, specific, informed and unambiguous...agreement." However, existing consent procedures result in large amounts of personal data processing without consumers understanding what they have "agreed" to, demonstrating that GDPR has not yet been enforced successfully against large digital platforms.[6] The reason for this delay, or failure, is primarily because of lack of enforcement in Ireland, the headquarters of both Google and Meta.[7] Therefore, in the baseline model we assume that the platform can always obtain consent from naive consumers, and we later examine the effects of enforcing informed consent in Section 6.2.

Lastly, the DMA requires that gatekeepers offer a less personalized version of their existing free platform service, also at no charge to the user (Recital 37). The less personalised alternative is required to not be different or of degraded quality compared to the service provided to the end users who provide consent (Section 6.1 provides a rationale for such requirement). The law is not specific about the characteristics of the less personalised option; one goal of our model is to evaluate the welfare effects of this "regulatory default option".

Before the passage of the DMA and DSA, the lack of enforcement in Ireland allowed gatekeepers to continue to process personal data in every member state. The European Parliament responded to the situation with the laws above, and designed them to be enforced by the European Commission, not member states.[8] The debate over what constitutes compliant data processing under the DMA has intensified since compliance was required in March of 2024, and will likely take time to be resolved as the enforcement process rolls out. Digital platforms must determine how to respond to the new laws, and a significant choice is the design of a menu of service versions that vary by the amount of data sharing. The impact of these laws on the services available to consumers is one reason that an economic model of key aspects of privacy and menu design is helpful to policymakers.

[6] Though many small businesses are complying at some cost (Johnson, 2024).

[7] A company's GDPR enforcer is the Data Protection Authority of the member state in which the company is headquartered. The Irish DPA is part of the Irish national strategy of appealing to companies on the basis that they will undercut other EU jurisdictions' attempts to regulate their businesses (Seabrooke and Wigan, 2014). For example, Ireland offers lower corporate tax rates, more lenient financial regulation, and also glacially slow and lenient data protection to its corporations. For more on Ireland's tax and regulatory environment, see J. Nader, *Fund finance in Ireland and Luxembourg: a comparative analysis*, Ogier, `https://www.ogier.com/news-and-insights/insights/fund-finance-in-ireland-and-luxembourg-a-comparative-analysis` (2022). On data protection, see *EU Court: Irish DPB must investigate nyob complaint*, `https://noyb.eu/en/eu-court-irish-dpc-must-investigate-noyb-complaint` (2025).

[8] The DMA's obligations came into force against designated gatekeepers only at the beginning of 2024, and non-compliance investigations that take one year to complete ended in spring 2025, both resulting in a finding of non-compliance. However, during the analysis year, Meta changed its terms, necessitating a second round of analysis.

1.2 Application: Meta's Menu

To make the model more concrete we offer a specific example, the case of Meta's menu for its network Facebook. Since its inception, Facebook offered the standard version of its personal social network to end users without any monetary charge. Instead, end users "agree" to share their personal data with Facebook when they create an account.[9] Over time the amount of personal data Facebook collected from its users—and used to target advertising—has grown. In particular, Facebook began to track its users as they traveled the open web using cookies and other tools, and used that information to target advertising and content when the user returned to Facebook. Because some users and civil society organizations became concerned about this tracking (for example, the Cambridge Analytica scandal[10]), over the decade from 2014-2024, Facebook gradually expanded the ability of users to modify what personal data was available to other people on Facebook, used by third parties, and used by Facebook and its services. In 2022, Facebook centralized many (though not all) of these settings in the "Privacy Center" where users could go to make specific choices about how much data to share and tracking to allow.[11] Therefore, if the user takes action, they can use a free version of the service that shares less personal data.

Facebook offered a third option in Europe with no advertising at all at a cost of 10 Euros per month, later lowered to 6 Euros per month.[12] The result of this evolution is a European menu for this popular platform in which a user can choose from a default that has high shared data and zero monetary charge, a second option with lower shared data and zero monetary charge, and a third option requiring payment of a monthly fee for access to the regular version of the service with no advertising. Our model solves for the monopolist's optimal menu of versions in a setting with institutional details similar to the EU and the design tools we see being used in the marketplace. The results of the model shed light on the menus that platforms offer and help regulators evaluate the welfare consequences of those menus.

1.3 Contribution to the Literature

Our work contributes to the extensive literature on the economics of privacy (see Acquisti et al., 2016 for an excellent overview). Privacy protection is a particularly salient concern in the digital economy, where ad-supported platforms collect vast amounts of consumer data, and a large empirical literature documents that consumers can be harmed by data processing (Grande et al., 2020; Aridor et al., 2023).

Recent works by Gradwohl (2018), Eilat et al. (2021), and Krähmer and Strausz (2023) study

[9] As of October 2025, users are not explicitly asked to consent to any data sharing during the account creation process. The page does contain a message that reads, "By clicking Sign Up, you agree to our Terms, Privacy Policy and Cookies Policy," with links to each of those agreements. See "Sign Up for Facebook," Facebook, `https://www.facebook.com/r.php?entry_point=login`.

[10] See, for example, Harbath, Katie, and Collier Fernekes, "History of the Cambridge Analytica Controversy," *Bipartisan Policy Center*, 16 Mar. 2023, `https://bipartisanpolicy.org/blog/cambridge-analytica-controversy/`.

[11] Introducing Privacy Center, `https://about.fb.com/news/2022/01/introducing-privacy-center/` (2022).

[12] Facebook and Instagram to offer subscription for no ads in Europe, Meta, `https://about.fb.com/news/2024/11/facebook-and-instagram-to-offer-subscription-for-no-ads-in-europe/` (2024).

menu pricing under the assumption that some consumers incur privacy losses from revealing information about their type, and characterize privacy-constrained mechanisms.[13] In our model, as in the canonical framework of Mussa and Rosen (1978), consumer type measures valuation for quality, but here quality embeds privacy losses, so net quality can become negative when the platform collects a sufficiently large amount of data. We show that when the platform can subsidize consumer participation (i.e., charge negative access prices), the optimal menu features bunching of intermediate types at an option with zero net quality and zero price, resembling the solution in models with countervailing incentives (Lewis and Sappington, 1989; Jullien, 2000).

We are not the first to study the impact of regulatory constraints on a monopolist's menu design. Besanko et al. (1987, 1988) show that introducing minimum quality standards in the canonical menu-pricing model induces bunching at the regulated quality for relatively low-valuation types—thereby alleviating the under-provision of quality for these types—but also causes the exclusion of the lowest ones. In our model, a strict hard cap on data trade may have similar effects, whereas default regulation entails different welfare trade-offs. The closest work to ours in this regard is Dworczak and Muir (2025), who study the design of outside options, with similar implications to our regulatory default, but focus on property rights, investment, and hold-up.

We show that the desirability and optimal design of regulation crucially depend on the presence and extent of consumer naiveté. Introducing naive consumers in our setting is both realistic and consistent with empirical evidence. In this respect, our work also contributes to the literature on second-degree price discrimination with behavioral consumers (see Kőszegi, 2014 for a survey, and Eliaz and Spiegler, 2006, 2008; Heidhues and Kőszegi, 2017; Johnen, 2019).

Finally, the privacy losses we model can also be interpreted as a reduced form for instrumental privacy preferences arising from consumers' anticipation that data collected by the platform may be used for targeted advertising or price discrimination (e.g., Ichihashi, 2020; Kirpalani and Philippon, 2020; Anderson et al., 2023; Bergemann and Bonatti, 2024; Galperti et al., 2024).[14] However, evidence in Lin (2022) suggests that both intrinsic and instrumental privacy preferences are empirically relevant.

1.4 Organization

We describe the model in Section 2 and the optimal menu of service levels in Section 3. We then introduce behavioral (naive) consumers in Section 4, before turning to two regulatory interventions—a *hard cap* on data trade and a regulated *default version*—in Section 5. We extend the model to endogenous service quality and to informed consent requirements in Section 6, before concluding in Section 7. All proofs are in Appendix A. The extracts of the GDPR, DMA, and DSA mentioned above are reproduced in Appendix B for the interested reader.

[13] The computer science literature (e.g., Pai and Roth, 2013)has instead used the concept of differential privacy.

[14] Our framework, however, does not capture information externalities among users who sell their data (Choi et al., 2019; Acemoglu et al., 2022; Bergemann et al., 2022). We view this omission as a minor limitation, since by 2025 the most valuable personal data for advertisers is highly individual and contextual rather than social (Sahni et al., 2019; Xiong et al., 2022).

2 Model with Sophisticated Consumers

A monopoly platform sells access to a service in exchange for payments and the ability to collect consumer data. We denote the amount of data collected by $D \geq 0$. Processing the collected data generates advertising revenues αD for the platform, i.e., better targeted ads are more profitable.[15]

Denoting by T the (possibly negative) price the consumer pays to access the platform service, the platform profit per user is given by

$$\alpha D + T.$$

The net quality of the platform service is defined as

$$q := v - L(D),$$

where $v > 0$ is the baseline service quality—the quality of a service that collects no data at all—and $L(\cdot)$ is a positive, strictly increasing and convex function that captures losses in privacy due to data collection.[16] For simplicity, we will assume that $L(D) := D^2/2$.

We describe our model with sophisticated consumers only, and introduce naive consumers in Section 4. A consumer's valuation for a service with quality level q and access price T is given by

$$\theta q - T,$$

where $\theta > 0$ is the consumer's private information. Every consumer has an outside option of zero.

The consumer's type θ is distributed over the interval $\Theta := [\underline{\theta}, \overline{\theta}]$ according to the distribution $F(\theta)$ with density $f(\theta) > 0$. We assume that both the virtual cost *à la* Baron and Myerson (1982) and the virtual value *à la* Myerson (1981) (henceforth "the virtual valuations") are everywhere positive and strictly increasing:

Assumption 1 (Monotone Virtual Valuations). *$\theta + F(\theta)/f(\theta)$ and $\theta - [1 - F(\theta)]/f(\theta)$ are increasing for all θ.*

By the Revelation Principle, the platform can offer a direct mechanism $\{D(\theta), T(\theta)\}_{\theta \in \Theta}$ specifying a data option and an access price as function of the consumer's reported type θ. Thus, we modify the standard menu-pricing model by allowing quality to become negative, i.e., $v < L(D)$ (hence $q < 0$) for sufficiently large D, and by letting the platform subsidize consumers (i.e., offer $T < 0$) to ensure participation when providing negative quality.[17]

[15] For tractability, we assume that the platform's marginal value of data, denoted by α, is constant across consumer types. Thus, we abstract from potential complementarities between data quality and user characteristics, allowing us to isolate the mechanism through which heterogeneous privacy preferences shape the optimal mechanism.

[16] Convexity of L suggests that the platform collects relatively harmless information first, and increasingly sensitive pieces as data collection expands. Each additional unit of data therefore entails a larger incremental loss of privacy.

[17] As argued in Monti et al. (2025), the platform can compensate consumers though complimentary ancillary services to avoid offering negative prices, which may expose the platform to arbitrage by bots (Bisceglia and Tirole, 2023).

3 Optimal Menus

3.1 Complete Information

Before analyzing the platform's screening problem, we characterize the solution under complete information, that is, when the platform knows each consumer's type θ.

For each type, the sum of the platform profit and the consumer's utility is

$$W(\theta) := \alpha D + \theta\left(v - \frac{1}{2}D^2\right). \tag{1}$$

Under complete information, the data options offered to consumers maximize social welfare, which is fully appropriated by the platform through the access price.

Proposition 1 (**Complete Information**). *Under complete information, the platform offers to each type θ the following data option and access price:*

$$D^{CI}(\theta) = \frac{\alpha}{\theta}, \quad T^{CI}(\theta) = \theta\left[v - \frac{1}{2}\left(\frac{\alpha}{\theta}\right)^2\right].$$

The complete-information payments are positive, $T^{\text{CI}}(\theta) > 0$, if and only if $\theta > \alpha/\sqrt{2v}$. Therefore, for all $\theta \in \left[\underline{\theta}, \alpha/\sqrt{2v}\right)$, under complete information, the platform subsidizes the consumer participation to compensate for the negative net quality. This negative-price region is empty if the baseline service quality is sufficiently large, so that $\underline{\theta} \geq \alpha/\sqrt{2v}$.

Intuitively, since a unit of data D has the same value to the platform across all consumers and data ranges, any variation in the efficient level of data collection ($D^{CI}(\theta)$, Figure 1a) is entirely driven by the different consumers' valuations for privacy. The corresponding prices ($T^{\text{CI}}(\theta)$, Figure 1b) are such that high-valuation consumers pay to keep their data and low-valuation consumers are paid to share their data.

3.2 Asymmetric Information

We now consider the platform's problem under asymmetric information. We begin with the information rent of a consumer who truthfully reveals their type θ, i.e., selects the option $(D(\theta), T(\theta))$ designed for them:

$$U(\theta) := \theta\left(v - \frac{1}{2}[D(\theta)]^2\right) - T(\theta). \tag{2}$$

A menu of contracts $\{D(\theta), T(\theta)\}_{\theta\in\Theta}$ is incentive compatible if and only if

$$U(\theta') = U(\theta) + \int_{\theta}^{\theta'} \left(v - \frac{1}{2}[D(\tilde{\theta})]^2\right) \mathrm{d}\tilde{\theta} \quad \forall \theta, \theta' \tag{IC}$$

and the following monotonicity condition holds:

$$D(\theta') \leq D(\theta) \quad \forall \theta' > \theta. \tag{M}$$

The individual-rationality constraints are

$$U(\theta) \geq 0 \quad \forall \theta. \tag{IR}$$

By the Revelation Principle, the platform's problem then writes as

$$\max_{\{D(\theta),U(\theta)\}_{\theta\in\Theta}} \int_{\underline{\theta}}^{\overline{\theta}} [W(\theta) - U(\theta)]\,\mathrm{d}F(\theta)$$
$$\text{s.t. (IR)-(IC)-(M).}$$

Differentiating (IC) implies that $U(\theta)$ is strictly decreasing (resp. increasing) for all $D(\theta) > \sqrt{2v}$ (resp. $D(\theta) < \sqrt{2v}$). Therefore, $U(\theta)$ attains its minimum value at all θ for which $D(\theta) = \sqrt{2v}$, i.e., for consumers obtaining zero net quality. The monotonicity condition (M) implies that there are two cutoffs, θ_1 and θ_2, with $\underline{\theta} \leq \theta_1 \leq \theta_2 \leq \overline{\theta}$ (at least one inequality being strict), such that $D(\theta) > \sqrt{2v}$ for $\theta \in [\underline{\theta}, \theta_1)$, $D(\theta) = \sqrt{2v}$ for $\theta \in [\theta_1, \theta_2]$, and $D(\theta) < \sqrt{2v}$ for $\theta \in (\theta_2, \overline{\theta}]$. We characterize the solution to the platform's problem in the following Proposition.

Proposition 2 (**Optimal Menu**).

1. *The platform offers the following continuous menu of data options*

$$D^*(\theta) = \begin{cases} \frac{\alpha}{\theta + \frac{F(\theta)}{f(\theta)}} & \text{for } \theta \in [\underline{\theta}, \theta_1), \\ \sqrt{2v} & \text{for } \theta \in [\theta_1, \theta_2], \\ \frac{\alpha}{\theta - \frac{1-F(\theta)}{f(\theta)}} & \text{for } \theta \in (\theta_2, \overline{\theta}]. \end{cases} \tag{3}$$

2. *The optimal prices $T^*(\theta)$ are negative and strictly increasing for $\theta \in [\underline{\theta}, \theta_1)$, zero for all $\theta \in [\theta_1, \theta_2]$, and positive and strictly increasing for $\theta \in (\theta_2, \overline{\theta}]$.*

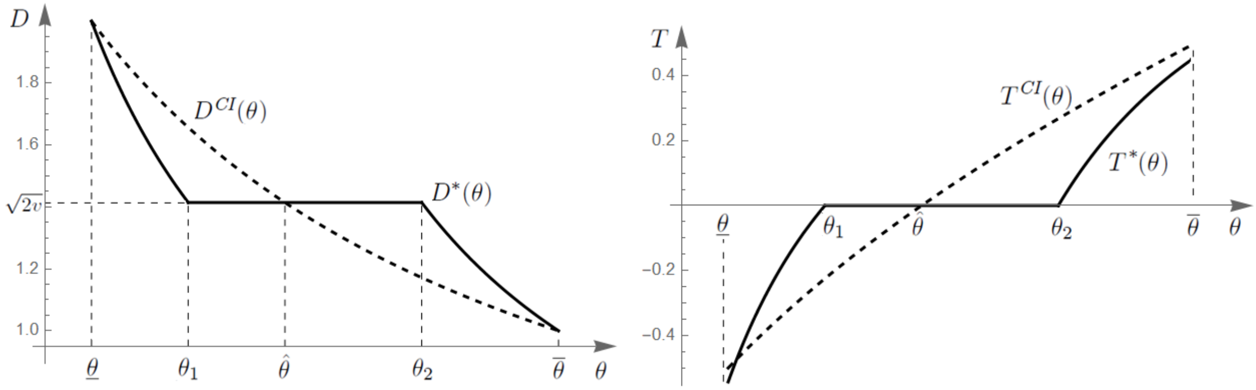
3. *If $\underline{\theta} < \alpha/\sqrt{2v} < \overline{\theta}$, the cutoffs $\underline{\theta} < \theta_1 < \theta_2 < \overline{\theta}$ are the unique solutions to*

$$\frac{\alpha}{\theta_1 + \frac{F(\theta_1)}{f(\theta_1)}} = \sqrt{2v} = \frac{\alpha}{\theta_2 - \frac{1-F(\theta_2)}{f(\theta_2)}}. \tag{4}$$

4. *If $\underline{\theta} \geq \alpha/\sqrt{2v}$, the interval $[\underline{\theta}, \theta_1)$ is empty; and if $\alpha/\sqrt{2v} \geq \overline{\theta}$, the interval $(\theta_2, \overline{\theta}]$ is empty.*

The platform-optimal data options and access prices are illustrated in Figure 1 for an example in which all the three regions in the optimal menu are non-empty.[18]

[18]Unless otherwise stated, all figures in the paper are computed for a uniform distribution of types with parameter values $\alpha = v = 1$, $\underline{\theta} = 0.5$, and $\overline{\theta} = 1$.



(a) Data options. (b) Prices.

Figure 1: Optimal menu under complete (dashed lines) vs asymmetric information (continuous lines), for $\alpha/\overline{\theta} < \sqrt{2v} < \alpha/\underline{\theta}$.

As noted in Section 2, defining quality as $q(\theta) := v - L(D(\theta))$, our model coincides with the standard monopoly screening problem along price and quality dimensions, except that we do allow for negative quality ($q(\theta) < 0$). This occurs in the case of intense data collection and implies the necessity of negative prices to satisfy consumers' participation constraint. Proposition 2 shows that the second-best menu mirrors that of a screening problem with *countervailing incentives*.[19]

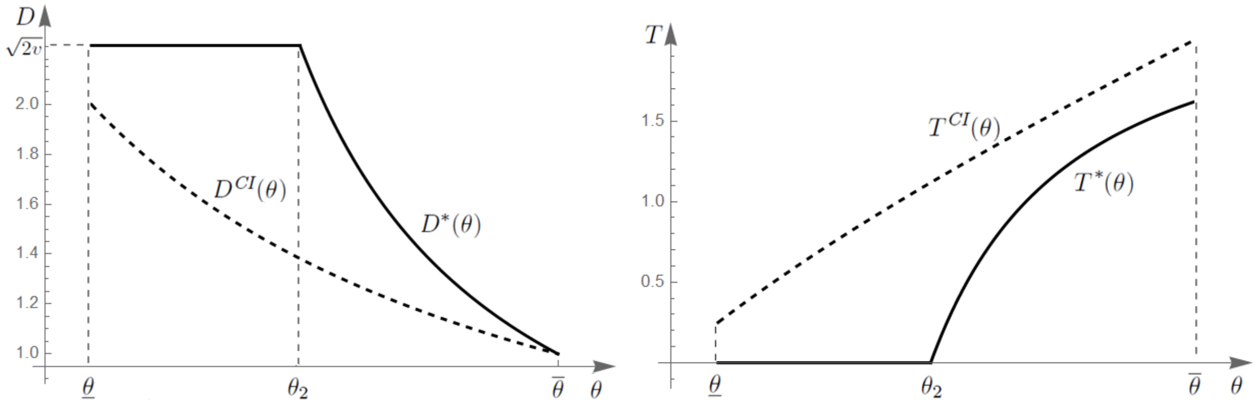
In particular, high consumer types $\theta > \theta_2$ obtain positive quality $q(\theta) > 0$ and pay $T(\theta) > 0$: they buy privacy. These types are tempted to understate their value for privacy, i.e., to *mimic* a lower type, to buy privacy at a lower price, as in quality-pricing models (Mussa and Rosen, 1978). To deter this behavior and limit these consumers' rents, the platform distorts the quality of the good being traded (i.e., privacy) downward relative to the first-best allocation. This distortion leads to inefficiently high data trade, $D^*(\theta) > D^{\text{CI}}(\theta)$ for all $\theta \in (\theta_2, \overline{\theta})$.

Conversely, low types $\theta < \theta_1$ receive negative quality $q(\theta) < 0$ and a positive transfer from the platform $T(\theta) < 0$, i.e., they sell personal data. Since such consumers are effectively *selling* quality (rather than buying it), they have incentives to overstate their cost of selling—to *mimic* a higher type—to obtain a higher transfer for their personal data, similar to the incentive to overstate costs in procurement models (Baron and Myerson, 1982). To deter this behavior and limit these consumers' rents, the platform distorts data collection downward, which results in inefficiently high privacy levels, $D^*(\theta) < D^{\text{CI}}(\theta)$ for all $\theta \in (\underline{\theta}, \theta_1)$.

The optimal mechanism bunches a positive-measure set of intermediate types into a *no-trade* option, characterized by zero net quality and zero transfer, $D(\theta) = \sqrt{2v}$ and $T(\theta) = 0$. Bunching is a consequence of the downward and upward pressures on data collection when trading under asymmetric information, resulting in a zero-trade, zero-price region. This bunching option features inefficiently low (resp. high) trade for relatively low (resp. high) types: letting $\hat{\theta} := \alpha/\sqrt{2v}$, consumers $\theta \in [\theta_1, \hat{\theta})$ (resp. $\theta \in (\hat{\theta}, \theta_2]$) trade too little (resp. too much) data.

[19] See Lewis and Sappington (1989), Maggi and Rodríguez-Clare (1995), Jullien (2000) and, more recently, Loertscher and Muir (2025) and Dworczak and Muir (2025).

Finally, when v is so large that $\underline{\theta} \geq \alpha/\sqrt{2v}$, buying data above the bunching level $D = \sqrt{2v}$ (which is socially excessive in this case) is too expensive for the platform. As a result, no consumers receive positive payments, although all consumers trade too much data: see Figure 2.



(a) Data options. (b) Prices.

Figure 2: Optimal menu under complete (dashed lines) vs asymmetric information (continuous lines), for $\sqrt{2v} = \sqrt{5} > \alpha/\underline{\theta}$.

The bunching region in Figure 2 is consistent with the business model of many digital platforms that offer "free" services. Products such as browsers, search engines, and social media often grant access to consumers without any monetary charge while engaging in personal data collection from which they profit (e.g., targeted advertising, AI training, etc). Some platforms offer both a free service and paid options that feature reduced data collection—e.g., Apple iCloud+, Mozilla Firefox Relay Premium. When the baseline quality level is high ($\sqrt{2v} \geq \alpha/\underline{\theta}$), our model provides an explanation for the lack of negative-price options on all these platforms.

However, one feature of our motivating example—Meta's pay-or-consent model—cannot be rationalized by this approach. For many years, Meta has offered two versions of its service with zero monetary charge but with substantially different privacy levels. Yet, almost all consumers permitted greater processing of their personal data by choosing the less private option. The next section shows how augmenting the model to include a fraction of naive consumers rationalizes such platform behavior.

4 The Role of Naive Consumers

4.1 Assumptions and Benchmarks

We now introduce naive consumers who are not responsive to the true quality of the platform service. Formally, denoting by $\xi \in \{0, 1\}$ a consumer's *behavioral type*, with $\xi = 0$ (resp. $\xi = 1$) if the consumer is *naive* (resp. sophisticated), their net utility from accessing the platform (relative to the outside option) is

$$\xi\theta\,[v - L(D)] - T.$$

From a behavioral perspective, naive consumers act as a sophisticated consumer with type $\theta = 0$. When computing welfare, however, the payoffs of a type-θ naive consumer coincide with those of a sophisticated consumer.

A fraction $\gamma \in (0, 1)$ consumers is naive, and the entire type (θ, ξ) is the consumer's private information. We allow for correlation between a consumer's behavioral type and their true underlying valuation for quality by assuming that θ follows a distribution $F_N(\theta)$ for naive consumers, which may differ from the distribution $F(\theta)$ for sophisticated consumers.

In what follows, we posit that the amount of data collection is $D \in [0, \overline{D}]$, with $\overline{D}$ reflecting an exogenous technological limit to data collection, and assume

Assumption 2 (Data Exploitation)**.** $\overline{D} > \max\{\sqrt{2v}, \alpha/\underline{\theta}\}$.

This assumption ensures that the technological constraint is never binding for sophisticated consumers and that collecting $\overline{D}$ from any consumer is socially inefficient. Moreover, we slightly generalize the monotone virtual cost assumption:

Assumption 1' (Monotone Virtual Valuations with Naive Consumers)**.** $\theta + [F(\theta) + \gamma/(1-\gamma)]/f(\theta)$ *and* $\theta - [1 - F(\theta)]/f(\theta)$ *are increasing in* θ for all θ.

Under complete information, the platform would engage in *data exploitation* by offering naive consumers the option $(D, T) = (\overline{D}, 0)$—that is, it would fully extract their data without compensating them—while still offering the menu characterized in Proposition 1 to sophisticated consumers.[20] Therefore, while the complete information solution is socially efficient for sophisticated consumers, this is not the case for naive consumers.

Under asymmetric information, in order to screen the behavioral type, the platform offers an incentive-compatible option targeted to naive consumers in addition to a menu of contracts $\{D(\theta), T(\theta)\}_{\theta \in \Theta}$ designed for sophisticated consumers.

4.2 Optimal Menu with Naive Consumers

Since naive consumers are non-responsive to privacy differences, they select the lowest-priced available option, as long as it does not involve a strictly positive price. Recall that, in any incentive-compatible menu designed for sophisticated consumers, the lowest price is associated with the data option $D(\underline{\theta})$ designed for the lowest type. Anticipating the behavior of naive consumers, the platform adds to the menu the data-exploitative option $(\overline{D}, T(\underline{\theta}))$, which all naive consumers choose.

Provided that $\underline{\theta} < \alpha/\sqrt{2v}$, the lowest-price item in the menu of Proposition 2 entails a monetary subsidy for using the service (i.e., $T^*(\underline{\theta}) < 0$). Therefore, the platform would have to pay all naive consumers a transfer, even though they would be willing to share all their data for free, i.e., to

[20] Similarly, if the platform could observe the consumer's behavioral type or use choice architecture that impacts naive and sophisticated consumers differently (e.g., if naive consumers could not move away from the "free" default option, because they face large search costs or lack an understanding of privacy terms and informed consent is enforced), the platform would offer the menu characterized in Proposition 2 to sophisticated consumers, and add the exploitative option $(\overline{D}, 0)$ for naive consumers. However, as discussed above, this outcome is inconsistent with the widespread absence of negative-price options on digital platforms.

accept $(\overline{D}, 0)$. This observation implies that there are opportunity costs to the platform of offering to sophisticated consumers their optimal tariff. When naive consumers are present, the platform can profitably revise its menu to reduce (or even eliminate) the payments to sophisticated types.

Formally, from the definition of $U(\cdot)$ in (2), we have that

$$T(\underline{\theta}) = \underline{\theta}\left(v - \frac{1}{2}D(\underline{\theta})^2\right) - U(\underline{\theta}). \tag{5}$$

Therefore, the platform's problem writes as

$$\max_{\{D(\theta),U(\theta)\}_{\theta\in\Theta}} (1-\gamma)\int_{\underline{\theta}}^{\overline{\theta}} [W(\theta) - U(\theta)]\,\mathrm{d}F(\theta) + \gamma\left[\alpha\overline{D} + \underline{\theta}\left(v - \frac{1}{2}D(\underline{\theta})^2\right) - U(\underline{\theta})\right]$$
$$\text{s.t. (IR)-(IC)-(M)},$$

where the rents are computed using constraint (IC) defined in Section 3.2.

Because the second summand in the platform's objective depends on $D(\underline{\theta})$, when optimizing pointwise with respect to $D(\theta)$, the derivative has a downward jump at $\theta = \underline{\theta}$. In a separating menu, this discontinuity would imply $D(\underline{\theta}) < D(\theta)$ for θ in a right-neighborhood of $\underline{\theta}$, which would violate the monotonicity condition (M). Therefore, for all $\gamma > 0$, the solution must exhibit bunching over the interval $\theta \in [\underline{\theta}, \theta_0]$, for some $\underline{\theta} < \theta_0 \leq \theta_1$. Intuitively, for any fixed $D(\underline{\theta})$, such bunching minimizes the inefficient downward distortion of data trade with higher types $\theta > \underline{\theta}$ which is needed to satisfy (M). Moreover, to save on the rents for types $\theta \in [\underline{\theta}, \theta_0]$ who select this bunching option, the platform reduces the data collection from higher types $\theta \in (\theta_0, \theta_1)$, as explained in Section 3.2. The optimal menu in the presence of naive consumers is characterized as follows.

Proposition 3 (Optimal Menu with Naive Consumers).

1. *For any $\gamma \in [0, 1)$, the platform offers the following continuous menu of data options, intended for sophisticated consumers:*

$$D^*(\theta) = \begin{cases} \frac{\alpha}{\theta_0 + \frac{F(\theta_0)}{f(\theta_0)} + \frac{\gamma}{(1-\gamma)f(\theta_0)}} & \text{for } \theta \in [\underline{\theta}, \theta_0], \\ \frac{\alpha}{\theta + \frac{F(\theta)}{f(\theta)} + \frac{\gamma}{(1-\gamma)f(\theta)}} & \text{for } \theta \in (\theta_0, \theta_1), \\ \sqrt{2v} & \text{for } \theta \in [\theta_1, \theta_2], \\ \frac{\alpha}{\theta - \frac{1-F(\theta)}{f(\theta)}} & \text{for } \theta \in (\theta_2, \overline{\theta}]. \end{cases} \tag{6}$$

2. *The optimal prices $T^*(\theta)$ are negative and increasing for $\theta \in [\underline{\theta}, \theta_1)$ (strictly so for $\theta > \theta_0$), zero for all $\theta \in [\theta_1, \theta_2]$, and positive and strictly increasing for $\theta \in (\theta_2, \overline{\theta}]$.*

3. *There exists a threshold $\hat{\gamma} \in [0, 1)$ such that, for all $\gamma < \hat{\gamma}$, the cutoffs $\underline{\theta} < \theta_0 < \theta_1 < \theta_2$ are the unique solutions to*

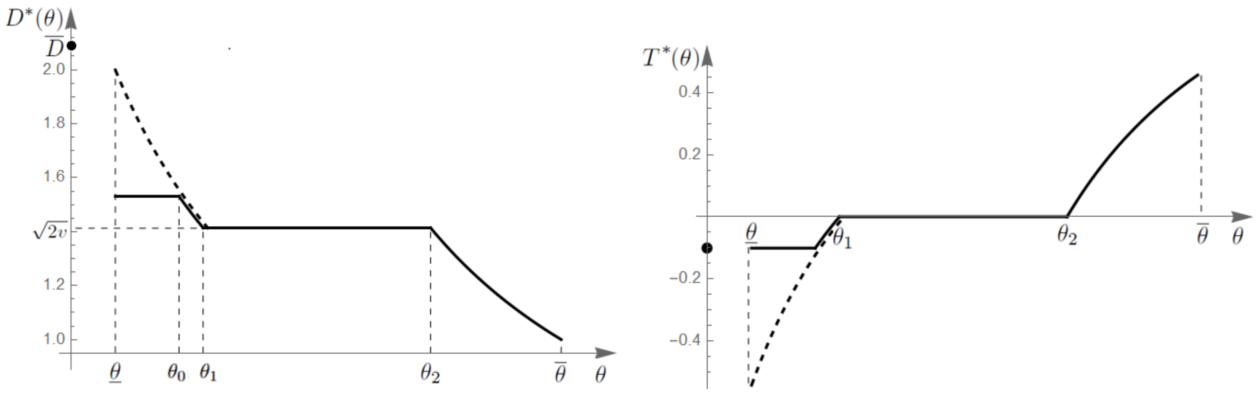
$$F(\theta_0) + \gamma\left(1 - F(\theta_0) - \frac{f(\theta_0)}{F(\theta_0)}\theta_0\right) = 0 \tag{7}$$

and

$$\frac{\alpha}{\theta_1 + \frac{F(\theta_1)}{f(\theta_1)} + \frac{\gamma}{(1-\gamma)f(\theta_1)}} = \sqrt{2v} = \frac{\alpha}{\theta_2 - \frac{1-F(\theta_2)}{f(\theta_2)}}. \tag{8}$$

4. *The threshold $\hat{\gamma}$ is decreasing in v and equals zero for all $\underline{\theta} \geq \alpha/\sqrt{2v}$. The interval $[\underline{\theta}, \theta_1)$ is empty (because $\theta_1 = \theta_0 = \underline{\theta}$) if $\gamma \geq \hat{\gamma}$. The interval $(\theta_2, \overline{\theta}]$ is empty if $\alpha/\sqrt{2v} \geq \overline{\theta}$.*

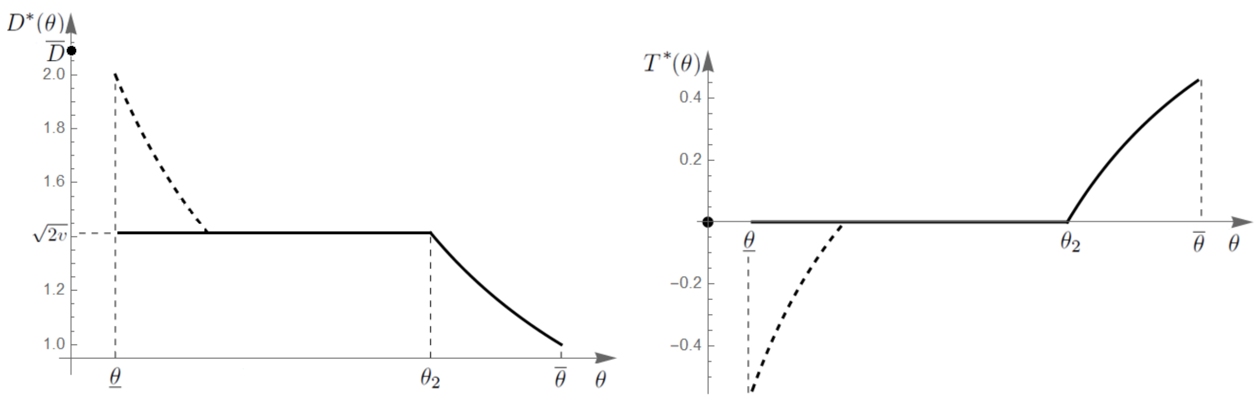
5. *The platform also offers the option $(\overline{D}, T^*(\underline{\theta}))$, intended for naive consumers.*



(a) Data options. (b) Prices.

Figure 3: Optimal menu for $\gamma = 0$ (dashed lines) vs $\gamma = 0.02 \in (0, \hat{\gamma})$ (solid lines).

Figure 3 illustrates the optimal menu. To reduce the payment made to the lowest sophisticated consumer type $\underline{\theta}$, which is also taken by all naive consumers, the platform reduces the amount of data collected from $\underline{\theta}$. Incentive compatibility then requires strictly reducing the data collected from any higher type that receives a positive transfer.



(a) Data options. (b) Prices.

Figure 4: Optimal menu for $\gamma = 0$ (dashed lines) vs $\gamma \geq \hat{\gamma}$ (solid lines).

These effects are stronger the larger the share of naive consumers in the market and, if there are sufficiently many naive consumers ($\gamma \geq \hat{\gamma}$), the platform forgoes any negative-price option, so as to avoid paying naive consumers for their data at all (see Figure 4). Unlike in the case of Proposition 2 with all sophisticated consumers, this can happen even if the baseline quality level is low ($\sqrt{2v} < \alpha/\underline{\theta}$). The following corollary summarizes the effects of naive consumers restricting attention to $\gamma < \hat{\gamma}$ (because the solution does not depend on γ when $\gamma \geq \hat{\gamma}$).

Corollary 1 (**Effect of Naive Consumers**). *Assume parameters are such that $\gamma < \hat{\gamma}$, and let θ_1 be the cutoff defined in equation* (4). *The share of naive consumers has the following effects:*

- *The data option $D^*(\theta)$ and the transfer $\left|T^*(\theta)\right|$ decrease with γ for all $\theta < \theta_1$.*
- *The rents of sophisticated consumers with $\theta < \theta_1$ and of naive consumers decrease strictly with γ.*

The platform's incentive to reduce payments to naive consumers leads to lower data collection and transfers for all sophisticated consumers with low privacy valuation—namely, the *data sellers*, types $\theta < \theta_1$ in a world without naive consumers (Proposition 2). In addition to exploiting consumer naiveté, the platform further exacerbates the inefficiently low level of data trade with these consumers and reduces their rents.

Conversely, the presence of naive consumers does not affect the platform's strategy *vis-à-vis* privacy-concerned, *data buyers*—types $\theta > \theta_2$. This is because naive consumers do not perceive any utility from quality and are therefore unwilling to pay positive prices. As a result, the options designed for high types remain unattractive to both consumers with low valuations for privacy and naive consumers.

As the share of naive consumers increases, each consumer is (at least weakly) worse off: sophisticated consumers who would be willing to trade data instead trade less and thus earn lower rents, while naive consumers receive lower transfers from the platform. Conversely, provided that $\overline{D}$ is sufficiently large, the platform finds data exploitation of naive consumers more profitable than data trade with sophisticated consumers. As a result, the platform profit increases with the fraction of naive consumers.

Our framework thus formally rationalizes Facebook's *pay-or-consent* model. The free default option, involving substantial data collection, corresponds to the contract $(D, T) = (\overline{D}, 0)$ designed for naive consumers. The more private option—also available for free, albeit only to users who are willing and able to locate it—corresponds to the bunching contract $(\sqrt{2v}, 0)$ intended for sophisticated consumers with relatively low privacy concerns. Finally, the more private subscription options correspond to the positive-price contracts offered to highly privacy-concerned sophisticated consumers in our model. The coexistence of two zero-price options with different privacy levels and the absence of negative-price options are thus consistent with a relatively large share of naive consumers in the market.

5 Privacy Regulation

The previous analysis has uncovered two sources of inefficiency: the platform exploits consumer naiveté by steering consumers who do not understand privacy policies into a data-exploitative option; and it distorts sophisticated consumers' privacy options relative to the socially efficient levels to limit rents. This section explores regulatory measures that mitigate both inefficiencies.

We study two types of regulatory interventions that shape the menus offered by platforms when user privacy is part of service quality: (1) imposing hard caps on data collection (Section 5.1), and (2) mandating privacy-enhancing default options (Section 5.2). As discussed in the Introduction (Section 1.1), these interventions correspond to measures in EU data protection law, and our model clarifies their implications for market outcomes and welfare. Finally, in Section 5.3 we characterize the consumer- and socially optimal combination of hard caps and default options.

5.1 Hard Cap on Data Trade

This section considers an explicit limit on the amount of data that can be collected from any consumer. This regulation aligns with the provisions in the EU's Digital Services Act which prohibit the trade of sensitive categories of personal data, such as health status or political beliefs, regardless of user consent and of the corresponding payments.

A hard cap on data collection at $D_1 < \overline{D}$ prevents the platform from exploiting naive consumers by offering contracts that extract all data—including the data above the level they would choose if they became informed—but might also bind *vis-à-vis* sophisticated consumers.

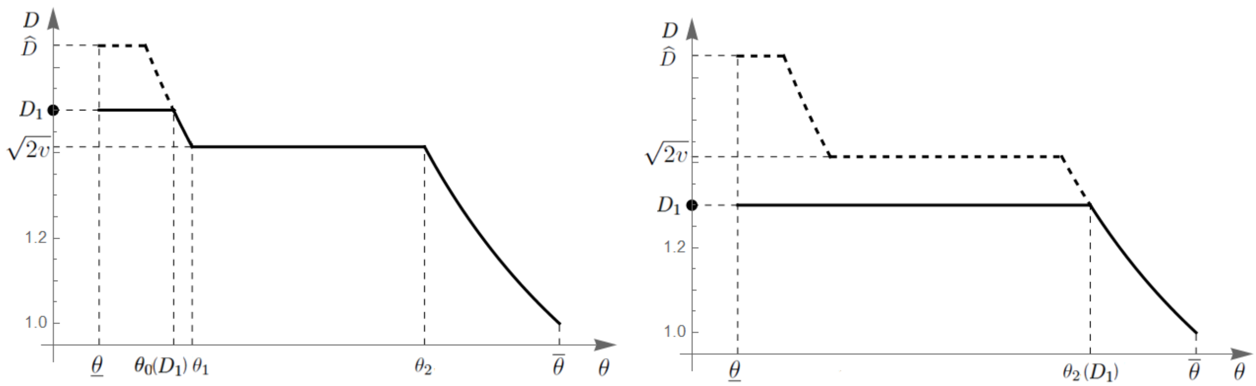
Proposition 4 (Optimal Menu with a Hard Cap). *Let $D^*(\theta)$ denote the platform-optimal menu in* (6)*. For all $D_1 < \overline{D}$, the platform offers the continuous menu of data options $D(\theta) = \min\{D_1, D^*(\theta)\}$ to sophisticated consumers and D_1 to naive consumers.*

Since the data options offered to sophisticated consumers are bounded away from $\overline{D}$, whether the cap also binds for some sophisticated type θ depends on how D_1 compares to the policy $D^*(\theta)$ in (6). Denote the bunching option offered to the low types as

$$\widehat{D} := D^*(\theta|\theta \leq \theta_0). \tag{9}$$

Because $D^*(\theta)$ is decreasing, if $D_1 \in (\widehat{D}, \overline{D})$, the menu offered to sophisticated consumers remains as in Proposition 3, and the hard cap binds only for naive consumers, who are offered the option $(D_1, T^*(\underline{\theta}))$. Such a hard cap increases both social and consumer welfare in the naive segment without affecting sophisticated consumers.

Any lower value of the hard cap, $D_1 < \widehat{D}$, also binds for some sophisticated consumers, preventing the platform from collecting more than D_1 data from types for which $D^*(\theta) > D_1$. The platform then optimally sets $D = D_1$ for all such consumers, without further distorting the optimal menu given in (6). We must now distinguish two cases.



(a) $D_1 \in (\sqrt{2v}, \widehat{D})$. (b) $D_1 \in (\alpha/\overline{\theta}, \sqrt{2v})$.

Figure 5: Optimal menu under *laissez faire* (dashed lines) and hard-cap regulation (continuous lines).

If $D_1 \in [\sqrt{2v}, \widehat{D})$, as in the left panel of Figure 5, the hard cap binds for sophisticated consumers $\theta \in [\underline{\theta}, \theta_0(D_1)]$, where the cutoff $\theta_0(D_1)$ satisfies

$$\frac{\alpha}{\theta_0 + \frac{F(\theta_0)}{f(\theta_0)} + \frac{\gamma}{(1-\gamma)f(\theta_0)}} = D_1. \tag{10}$$

All these consumers and naive consumers are offered the option $(D_1, T^*(\theta_0(D_1)))$, where $T^*(\theta_0(D_1)) < 0$ because the platform cannot collect more than D_1 and naive consumers always choose the lowest-priced available option. Because transfers are capped at $T^*(\theta_0(D_1)) > T^*(\theta)$ for $\theta \in [\underline{\theta}, \theta_0(D_1)]$, the net effect of the hard cap on naive consumers' surplus is in general ambiguous—it depends on whether the gains from additional privacy outweigh the reduction in transfers from the platform. By contrast, the inability to trade more than D_1 units of data unambiguously lowers consumer surplus for sophisticated consumers and reduces social welfare in this segment, since the types affected by the hard-cap regulation already trade inefficiently little data under *laissez faire*.

A tighter cap $D_1 < \sqrt{2v}$, as in the right panel of Figure 5, binds for all sophisticated consumers $\theta \in [\underline{\theta}, \theta_2(D_1)]$, where θ_2 is defined in 8. In this case, the platform chooses between two strategies. As long as D_1 exceeds a threshold $\widehat{D}_1$, the regulatory hard cap becomes the free option, and it is chosen by all naive consumers and by sophisticated consumers $\theta \in [\underline{\theta}, \theta_2(D_1)]$. If, however, $D_1 < \widehat{D}_1$, the platform switches to a pure *subscription model*, i.e., it offers both D_1 and the more private options $D \in (\alpha/\overline{\theta}, D_1)$ at positive subscription prices, excluding naive consumers and possibly some low-θ sophisticated consumers. The threshold $\widehat{D}_1 \in (0, \sqrt{2v})$ is decreasing in γ, since a larger share of naive consumers makes it more costly to exclude them in order to let sophisticated consumers pay for the data option D_1. Similar to the effects of minimum quality standards in Besanko et al. (1987, 1988), such a stringent hard-cap regulation, while efficiently enhancing privacy for consumers with high privacy valuations, has ambiguous overall effects on both consumer and social welfare.

Summing up, while hard-cap regulation protects naive consumers from data exploitation, any cap that constrains also allocations *vis-à-vis* sophisticated consumers entails potential downsides.

The main limitation of hard-cap regulation is that it can enhance privacy for high-θ consumers only by inefficiently dampening data trade among those with low privacy valuations. *A fortiori*, hard-cap regulation cannot induce the platform to offer efficient negative-price data options when $\gamma \geq \hat{\gamma}$. The next section shows that regulating default options can be an effective policy instrument to induce this outcome.

5.2 Default Regulation

Absent regulation, if $\gamma \geq \hat{\gamma}$ the platform steers naive consumers toward an option that collects large amounts of data at zero price. Because naive consumers focus solely on prices and disregard privacy losses, they do not find it worthwhile to navigate the settings in search of a more private free alternative, and therefore select the exploitative free option. Regulation that simply requires the platform to set its existing, more private free option $(\sqrt{2v}, 0)$ as the *default option* might appear to eliminate the salience advantage of the exploitative option and to address the problem. Naive consumers, however, will still select the exploitative option if the platform compensates them for their switching cost, which we assume the platform can design to be negligible. Making $(\sqrt{2v}, 0)$ the default will then have no effect on the equilibrium menu.[21]

We now consider a stronger form of regulation that requires the platform to offer a free default option providing a more stringent level of privacy protection than the already existing more private version.[22] In particular, we require the platform to make available to consumers the version $(D, T) = (D_0, 0)$, with $D_0 < \sqrt{2v}$. This version provides a consumer of type θ with net utility

$$U_0(\theta) := \theta \left(v - \frac{1}{2} D_0^2 \right) > 0.$$

The regulatory option is akin to strengthening the consumers' property rights over their privacy. In particular, default regulation gives consumers the option to limit data collection in the absence of payments. As in the property rights model of Dworczak and Muir (2025), this option generates a strictly positive, type-dependent outside option $U_0(\theta)$, making the analysis analogous to a screening model with countervailing incentives.

In Proposition 5 below, we consider the effect of free default options D_0 in the range $(\alpha/\overline{\theta}, \sqrt{2v})$ that illustrates our richest results.[23]

[21] Because the regulator cannot control all aspects of the platform's choice architecture, the platform can use design features and nudges to bring consumers' switching costs close to zero (this can be done by repeatedly prompting users to change their default, marketing the superiority of the platform-preferred option, or displaying the choice screen at moments when switching is most likely). More generally, if naive consumers' switching costs are sufficiently small—or can be made small enough—the default has no effect on the equilibrium outcome.

[22] For example, in the case of Facebook, a stricter option could limit the use of personal data to the same day it was collected. If the default version offered the same data collection as the regular version, but with the addition of a time limit, it would be stricter.

[23] In particular, for $D_0 \leq \alpha/\overline{\theta}$, the platform would not offer positive-price options. Conversely, default regulation has no effect for $D_0 \geq \sqrt{2v}$, because in that case the regulatory option would be equal to or dominated by the free option $(\sqrt{2v}, 0)$ under *laissez faire*.

Proposition 5 (Optimal Menu under Default Regulation).

1. *For any free default option* $D_0 \in (\alpha/\overline{\theta}, \sqrt{2v})$*, the platform offers the following continuous menu of data options, intended for sophisticated consumers:*

$$D^R(\theta) = \begin{cases} \frac{\alpha}{\theta_0 + \frac{F(\theta_0)}{f(\theta_0)} + \frac{\gamma}{(1-\gamma)f(\theta_0)}} & \text{for } \theta \in [\underline{\theta}, \theta_0], \\ \frac{\alpha}{\theta + \frac{F(\theta)}{f(\theta)} + \frac{\gamma}{(1-\gamma)f(\theta)}} & \text{for } \theta \in (\theta_0, \theta_1(D_0)), \\ D_0 & \text{for } \theta \in [\theta_1(D_0), \theta_2(D_0)], \\ \frac{\alpha}{\theta - \frac{1-F(\theta)}{f(\theta)}} & \text{for } \theta \in (\theta_2(D_0), \overline{\theta}]. \end{cases} \tag{11}$$

2. *The optimal prices* $T^R(\theta)$ *are negative and increasing for* $\theta \in [\underline{\theta}, \theta_1(D_0))$*, zero for all* $\theta \in [\theta_1(D_0), \theta_2(D_0)]$*, and positive and strictly increasing for* $\theta \in (\theta_2(D_0), \overline{\theta}]$*.*

3. *There exists a decreasing threshold* $\hat{\gamma}(D_0) \in [0, 1)$ *such that, for all* $\gamma < \hat{\gamma}(D_0)$*, the cutoffs* $\underline{\theta} < \theta_0 < \theta_1(D_0) < \theta_2(D_0)$ *are the unique solutions to equation* (7) *and*

$$\frac{\alpha}{\theta_1 + \frac{F(\theta_1)}{f(\theta_1)} + \frac{\gamma}{(1-\gamma)f(\theta_1)}} = D_0 = \frac{\alpha}{\theta_2 - \frac{1-F(\theta_2)}{f(\theta_2)}}. \tag{12}$$

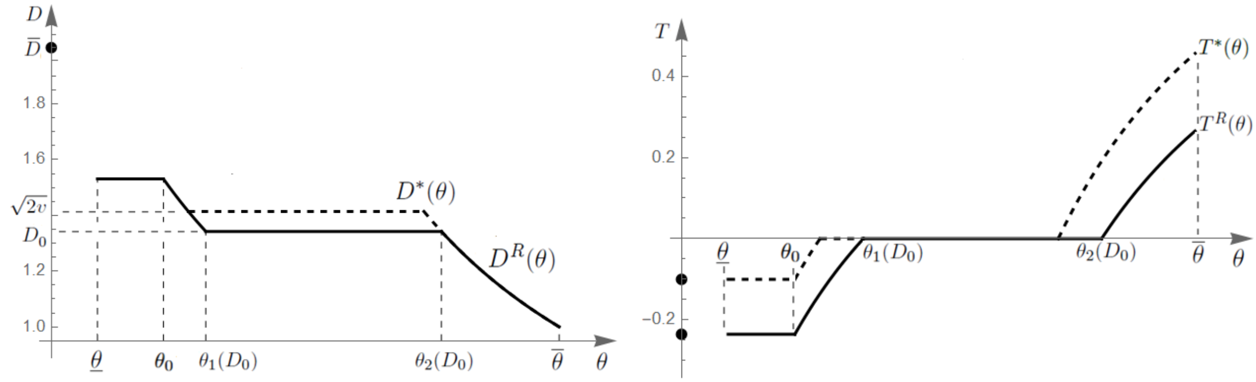
4. *The interval* $[\underline{\theta}, \theta_1(D_0))$ *is empty for all* $\gamma \geq \hat{\gamma}(D_0)$*.*

5. *The platform also offers the option* $(\overline{D}, T^R(\underline{\theta}))$*, intended for naive consumers.*

Default regulation therefore affects the data options only through the continuity condition of the optimal menu. Therefore, any consumer θ who does not select the free option $\sqrt{2v}$ under *laissez faire* and does not select the regulated default option D_0 chooses the same data option $D^*(\theta)$ in both cases. Now consider the set of types $\theta \in [\theta_1(D_0), \theta_2(D_0)]$ that select the default option. Denote by $\theta_i(\sqrt{2v})$ for $i = 1, 2$ the corresponding cutoffs in the *laissez-faire* outcome in Proposition 3. As the default option D_0 decreases from $D = \sqrt{2v}$, this region shifts to the right, as illustrated in Figure 6a.

These results also shed light on why privacy settings that platforms allow users to select for free are often difficult to find. In the case of Facebook, a user can opt out of being tracked around the open web and having personal data shared with third parties, for example. But these settings are not located all in one place, the choice locations are far down choice menus, and the labeling is not intuitive.[24] Such choice architecture requires a user to expend time and effort to select the more private, and still free, option, which lowers the likelihood that they do so. Our model justifies this design choice on the part of the platform by showing that a free, high-privacy alternative increases the incentive cost of collecting consumer data. The higher-privacy alternative improves outcomes for all consumers, who choose options with more privacy guarantees or better prices (consistent with Figure 6b and Corollary 2).

[24] For some depictions of the complexity of the process, see the appendix in Monti et al. (2025).



(a) Data options. (b) Prices.

Figure 6: Optimal menus under default regulation (continuous line) and *laissez faire* (dashed line), for $\gamma = 0.02 \in (0, \hat{\gamma}(\sqrt{2v}))$.

At the same time, the regulatory default option reduces platform profits. As such, a platform that has the choice between making such a version the default versus creating significant user search costs to locate and choose it will opt for the latter.

Corollary 2 (**Effects on Consumer Surplus**)**.** *For all* $\gamma \in [0, 1)$ *and* $D_0 \in (\alpha/\overline{\theta}, \sqrt{2v})$*, default regulation has the following effects on sophisticated consumers:*

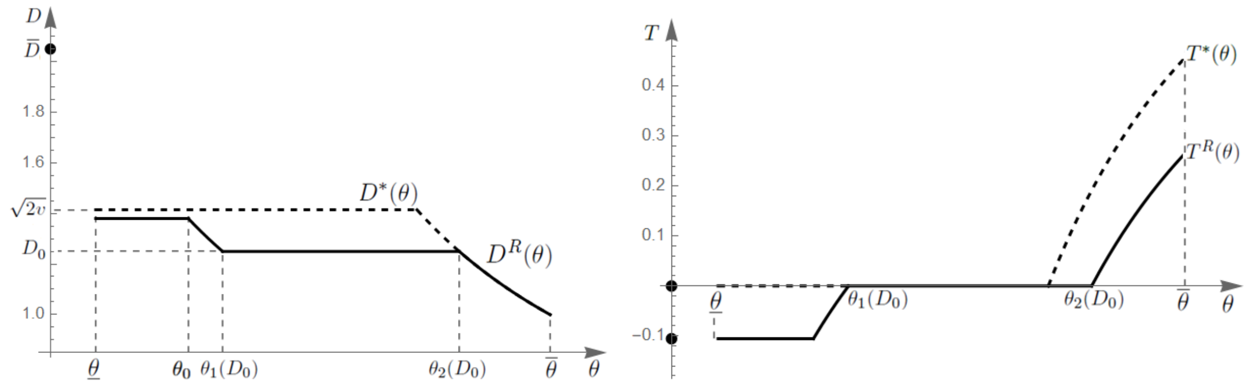
- *The privacy level for types* $\theta_1(\sqrt{2v}) < \theta < \theta_2(D_0)$ *strictly increases.*
- *The access price for types* $\theta \in [\underline{\theta}, \theta_1(D_0))$ *and* $\theta \in (\theta_2(\sqrt{2v}), \overline{\theta}]$ *strictly decreases.*
- *All consumers' rents strictly increase.*

While these results also hold if all consumers are sophisticated, default regulation may in addition have the following effects in the presence of (an intermediate fraction of) naive consumers.

Corollary 3 (**Data Market Creation and Payments for Data**)**.** *For any regulatory default* D_0 *such that* $\hat{\gamma}(\sqrt{2v}) < \gamma < \hat{\gamma}(D_0)$*:*

- (Data market creation.) *The platform introduces negative-price options with* $D^R(\theta) \in (D_0, \sqrt{2v})$ *for sophisticated consumers* $\theta \in [\underline{\theta}, \theta_1(D_0))$*.*
- (Compensation for data exploitation.) *The platform collects maximal data* $\overline{D}$ *from naive consumers but issues them a positive payment* $T^R(\underline{\theta}) < 0$*.*

As in Section 4.2, when naive consumers are sufficiently numerous, the platform refrains from trading additional data (i.e., trading $D > D_0$) with low-θ sophisticated consumers to avoid incurring the corresponding transfers to naive consumers. However, if the share of naive consumers is intermediate and the mandated default D_0 is sufficiently low, the platform finds it profitable to expand data trade with these low-θ sophisticated consumers, even though this entails making payments to naive consumers as well. This is illustrated in Figure 7.



(a) Data options.

(b) Prices.

Figure 7: Optimal menus under default regulation (continuous line) and *laissez faire* (dashed line), for $\gamma = 0.04 \in (\hat{\gamma}(\sqrt{2v}), \hat{\gamma}(D_0))$.

This *data market creation effect* of regulation alleviates the distortion toward excessive privacy provision that default regulation would otherwise impose on sophisticated consumers with low privacy valuations. Although default regulation is insufficient on its own to prevent the exploitation of naive consumers—and therefore does not increase social welfare in that segment—it still raises their surplus. Through the data market creation effect, default regulation ensures that naive consumers receive a transfer from the platform in exchange for their data.[25]

5.3 Optimal Regulation

The previous analysis implies that hard-cap and default regulations are complementary instruments: hard-cap regulation protects naive consumers from exploitation, while default regulation improves the efficiency of data collection for sophisticated consumers. Building on these insights, we characterize the optimal joint regulation of a hard cap and a free default.

Suppose, first, that the regulator follows a pure consumer surplus standard. As explained in Section 5.2, increasing the privacy level associated with the free regulatory default option D_0 enables consumers to obtain better privacy terms, better subscription prices, or both. Therefore, setting the lowest default option that ensures platform viability would be optimal. The optimal hard cap then trades off consumers' privacy gains against the lower monetary transfers they obtain under a lower D_1.

Consider, next, a regulator who maximizes social welfare, which is closer to the *proportionality* criteria adopted in the EU. Using the results in Propositions 4 and 5, the optimal regulation can

[25] For $\gamma < \hat{\gamma}(\sqrt{2v})$, default regulation generates a *data market expansion effect*, with similar consequences: additional sophisticated consumers are offered negative-price options, and naive consumers receive larger transfers. By contrast, if the fraction of naive consumers is too large relative to D_0 (i.e., $\gamma \geq \hat{\gamma}(D_0)$), the platform refrains from offering negative-price options even under default regulation, so naive consumers are unaffected.

be characterized as the solution to the following problem:

$$\max_{D_0 \le D_1} (1-\gamma) \int_{\underline{\theta}}^{\overline{\theta}} \Big[\alpha \min\{D_1, D^{\mathrm{R}}(\theta)\} + \theta\Big(v - \frac{1}{2}[\min\{D_1, D^{\mathrm{R}}(\theta)\}]^2\Big)\Big] \,\mathrm{d}F(\theta)$$
$$+ \gamma\Big[\alpha D_1 + \mathbb{E}_{F_N}[\theta]\Big(v - \frac{1}{2}D_1^2\Big)\Big], \tag{13}$$

where the policy $D^{\mathrm{R}}(\theta)$ denotes the monopolist's optimal menu in (11), which is itself a function of the regulatory default option D_0.

In Proposition 6 below, we characterize the welfare-optimal (D_0, D_1) under the following assumption on the distribution of consumer types, which we refer to as "uniform bounds."

Assumption 3 (Uniform Bounds)**.** *The distribution F has uniform bounds, for all $\theta \in [\underline{\theta}, \overline{\theta}]$,*

$$0 < m \le f(\theta) \le M < \infty, \tag{14}$$

$$f'(\theta) \le 0, \tag{15}$$

$$B := \sup_{\theta \in [0,1]} \frac{|f'(\theta)|}{f(\theta)^2} < \infty. \tag{16}$$

(These conditions are sufficient only, and the results in Proposition 6 can be shown numerically to hold for a wider class of distributions.)

Proposition 6 (**Welfare-Optimal Regulation**)**.** *Let F have uniform bounds with $M/m < 2 - B$ and assume v and $\mathbb{E}_{F_N}[\theta]$ are sufficiently large. Then, there exists a threshold $\gamma^* \in (0,1)$ such that, for all $\gamma < \gamma^*$, the welfare-optimal regulation is such that $D_1^W > D_0^W$.*

- *The optimal D_0^W satisfies*

$$\mathbb{E}_F\left[\theta \mid \theta \in [\theta_1(D_0^W), \theta_2(D_0^W)]\right] = \alpha / D_0^W, \tag{17}$$

  with $\theta_1(D_0)$ and $\theta_2(D_0)$ defined in (12)*.*

- *The optimal D_1^W solves*

$$(1-\gamma)F(\theta_0(D_1^W))\left(\alpha - \mathbb{E}_F[\theta|\theta \le \theta_0(D_1^W)]D_1^W\right) + \gamma\left(\alpha - \mathbb{E}_{F_N}[\theta]D_1^W\right) = 0, \tag{18}$$

  with $\theta_0(D_1)$ defined in (10)*.*

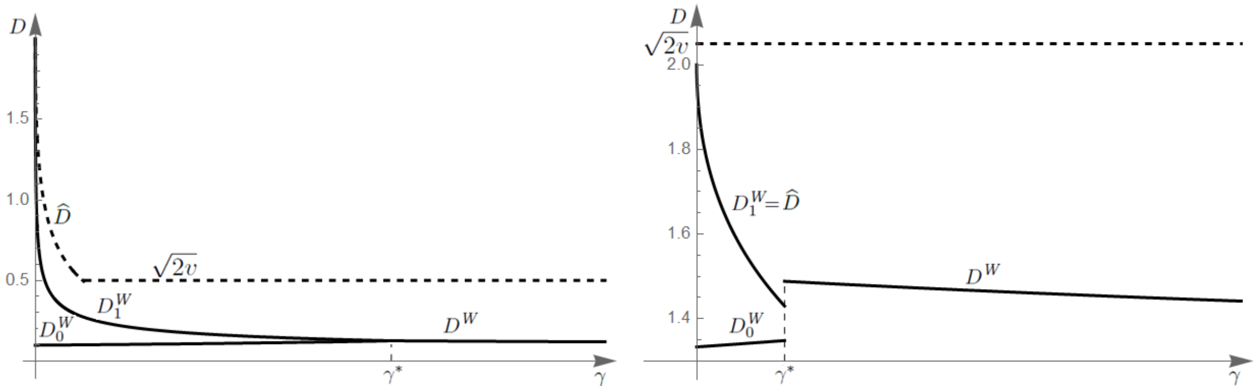
- *D_0^W is increasing in γ, and D_1^W is decreasing in γ.*

For all $\gamma \ge \gamma^*$, the welfare-optimal regulation is $D_0^W = D_1^W = D^W$, where D^W solves

$$(1-\gamma)F(\theta_2(D^W))\left(\alpha - \mathbb{E}_F[\theta|\theta \le \theta_2(D^W)]D^W\right) + \gamma\left(\alpha - \mathbb{E}_{F_N}[\theta]D^W\right) = 0 \tag{19}$$

and is decreasing in γ.

Proposition 6 restricts attention to parameters such that the welfare-optimal regulation is binding for sophisticated consumers, i.e., $D_0^{\mathrm{W}} < \sqrt{2v}$, and $D_1^{\mathrm{W}} < \widehat{D}$ (with $\widehat{D}$ defined in (9)), whenever $D_0^{\mathrm{W}} < D_1^{\mathrm{W}}$ (see Figure 8a). This requires, respectively, that v and $\mathbb{E}_{F_N}[\theta]$ are sufficiently large—the proof of Proposition 6 provides explicit conditions. If these conditions were violated, the optimal hard cap would be $D_1^{\mathrm{W}} = \widehat{D}$ (see Figure 8b), and the optimal default option would be $D_0^{\mathrm{W}} = \sqrt{2v}$, i.e., the value under *laissez faire*.



(a) $D_1^{\mathrm{W}} < \widehat{D}$ (for $\sqrt{2v} = 0.5$ and $\overline{\theta} = 20$). (b) $D_1^{\mathrm{W}} = \widehat{D}$ (for $\sqrt{2v} = 2.05$ and $\overline{\theta} = 1$).

Figure 8: Optimal default and hard-cap regulation, with $F_N = F \sim \mathcal{U}[0.5, \overline{\theta}]$.

If γ is sufficiently small, the platform sets a relatively private default option, which entails data market creation if v is large enough, and at the same time prevents the data exploitation of naive consumers through the hard cap, which is binding also for sophisticated consumers with low valuations for privacy, $\theta \in [\underline{\theta}, \theta_0(D_1^{\mathrm{W}})]$, if the average naive type $\mathbb{E}_{F_N}[\theta]$ is sufficiently large.

The welfare-optimal free default option D_0^{W} follows a simple principle: it coincides with the welfare-optimal pooling allocation for the set of types that choose it in equilibrium, $[\theta_1(D_0^{\mathrm{W}}), \theta_2(D_0^{\mathrm{W}})]$. This is because D_0 only determines the position of the bunching region (through the continuity of the optimal menu), without affecting the privacy policies designed for types outside this region.

If all consumers are sophisticated, the welfare-optimal default option induces inefficiently low privacy provision for low types in order to efficiently enhance privacy for high types. As a result, absent the need to protect naive consumers from data exploitation, a hard cap on data trade is either irrelevant (if D_1 is too high) or harmful to social welfare. Conversely, when naive consumers are present, imposing a hard cap allows the regulator to *de facto* choose the data option that will be offered to naive consumers, protecting them from data exploitation. Therefore, the welfare-optimal hard cap is always binding for naive consumers and may also bind for low sophisticated types.

As the fraction of naive consumers increases, the regulator becomes increasingly concerned with two issues: (1) excessive data trade by naive consumers, implying that D_1^{W} decreases with γ, and (2) over-provision of privacy for low-θ sophisticated consumers—because their data options are inefficiently distorted downward by the platform to limit payments to naive consumers—implying that D_0^{W} increases with γ. Consequently, for sufficiently large γ, the constraint $D_1 \geq D_0$ binds, and the regulator optimally sets a default option equal to the hard cap: $D_0^{\mathrm{W}} = D_1^{\mathrm{W}} = D^{\mathrm{W}}$. Since this default is offered to all naive consumers—whose optimal pooling allocation is $\alpha/\mathbb{E}_{F_N}[\theta]$—and to sophisticated consumers with low privacy concerns ($\theta \leq \theta_2(D^{\mathrm{W}})$)—whose optimal pooling allocation is strictly higher under our assumptions—it follows that D^{W} decreases with the market share γ of naive consumers.

6 Extensions and Robustness

The previous analysis has assumed that (i) the baseline service quality v is exogenous and set at the same level for all consumers, and (ii) naive consumers can choose any option without understanding the associated privacy terms. This section extends our model to endogenous service quality (Section 6.1) and to the enforcement of informed-consent procedures (Section 6.2).

6.1 Endogenous Service Quality

In this section, we endogenize service quality and study how regulation affects the platform's incentives to provide high-quality services. Suppose that the platform can offer personalized quality levels, i.e., a menu $(v(\theta), D(\theta), T(\theta))_{\theta\in\Theta}$. Quality provision entails a fixed investment cost $\Psi(\overline{v})$, where $\Psi(\cdot)$ is an increasing and convex function satisfying the usual assumptions $\Psi(0) = \Psi'(0) = 0$ and $\Psi(\infty) = \infty$, and $\overline{v} := \max_{\theta\in\Theta} v(\theta)$. That is, the platform incurs an investment cost corresponding to the "full-quality service" it decides to offer, whereas—since it provides digital services—the marginal cost of quality provision is zero.

Yet, for screening reasons, the platform could in principle choose to under-provide quality to some consumers, i.e., to offer $v(\theta) < \overline{v}$ for some types θ.[26] However, the consumers' rents—and thus the incentive-compatibility constraints—depend on the net quality $q(\theta) = v(\theta) - L(D(\theta))$ only. Hence, for any given $q(\theta)$, reducing $v(\theta)$ below the highest service quality offered to the other types is dominated by holding quality high (which entails no marginal cost) and increasing data-collection revenues. We formalize this intuition in the following lemma.

Lemma 1 (**Screening on Service Quality**). *Absent regulation, the platform finds it optimal to offer the same baseline service quality to all consumers.*

We now turn to default regulation. An important observation is that, if the regulator is silent about service quality, there is an obvious way for the platform to circumvent default regulation: it can degrade the service quality associated with the regulatory option to make it unappealing to any consumer. An example is Facebook's response to the adoption of the DMA, when the platform introduced a version of its service that used less personal data for ad targeting but showed full-screen, unskippable ads, a quality degradation.[27] The DMA requires that a gatekeeper offer a

[26] To avoid corner solutions, we assume that quality provision costs are large enough such that the technological constraint on data collection does not bind for sophisticated consumers, i.e., as in the baseline model, $D(\theta) < \overline{D}\ \forall\theta$.

[27] Facebook and Instagram to offer subscription for no ads in Europe, Meta, `https://about.fb.com/news/2024/11/facebook-and-instagram-to-offer-subscription-for-no-ads-in-europe/` (2024).

less personalized version of the "same service" exactly to guard against quality degradation that effectively coerces users into consuming a less private version in order to retain access to the regular service. In our model, this strategy corresponds to setting a service quality level $v_0 < D_0^2/2$, causing the regulatory default to become unappealing to any consumer and allowing the platform to continue offering its *laissez-faire* menu.

Lemma 2 (**Quality Degradation**). *Default regulation is inconsequential if the platform can degrade the service quality of the regulatory default option, i.e., offer* $(v_0, D_0, 0)$ *with* $v_0 < D_0^2/2$.

To prevent this behavior, the regulator must specify that the option $(D_0, 0)$ shall come at the full service quality, i.e., $v_0 = \overline{v}$. Under this additional requirement, both with and without default regulation, the platform offers the same quality level $\overline{v}$ to all consumers.[28] Nonetheless, default regulation alters the platform's incentives regarding service quality provision.

Proposition 7 (**Optimal Service Quality under Default Regulation**). *For any regulatory default option* D_0 *with a "full service quality" requirement* ($v_0 = \overline{v}$)*, the platform sets* $\overline{v}^R = D_0^2/2$. *Therefore, more stringent default regulation lowers baseline service quality.*

The unregulated platform sets the optimal baseline quality level, $\overline{v}^*$, taking into account that it will be able to collect $D = \sqrt{2\overline{v}^*}$ for free. When regulation constrains it to collect only $D_0 < \sqrt{2\overline{v}^*}$ for free, the platform invests less in service quality. In particular, it is always optimal for the platform to set service quality so that consumers taking the zero-price option obtain no rent.[29]

Suppose now that the regulator has commitment power, i.e., can commit to D_0 before the platform sets its service quality. The first-order condition for a binding welfare-optimal regulatory default option is given by

$$(1-\gamma)\int_{\theta_1(D_0)}^{\theta_2(D_0)} (\alpha - \theta D_0)\, \mathrm{d}F(\theta) + \left[(1-\gamma)\mathbb{E}_F[\theta] + \gamma \mathbb{E}_{F_N}[\theta] - \Psi'\left(\frac{1}{2}D_0^2\right)\right] D_0 = 0, \tag{20}$$

from which it follows that, relative to the optimal D_0^{W} with exogenous quality in (17)—which maximizes the first term in equation (20)—the regulator exploits commitment power to correct inefficiencies in quality provision.

Proposition 8 (**Optimal Default Regulation with Endogenous Service Quality**). *When the platform can adjust its service quality in response to regulation, the welfare-optimal default option is higher than* D_0^W *in* (17) *if and only if the platform would provide inefficiently low quality at* $D = D_0^W$.

[28] Indeed, regulation gives the platform even stronger incentives to maintain uniform quality, since setting $v(\theta) < \max_{\theta' \neq \theta} v(\theta')$ for some θ may also require offering either more privacy or a lower transfer to induce that consumer away from the regulatory option, which by regulation is associated with higher baseline quality.

[29] As a result, unlike in the model with exogenous baseline service quality, consumers $\theta \in [\theta_2(\sqrt{2v}), \theta_2(D_0))$ who select the more private regulatory default option rather than the option $D^*(\theta) \in (D_0, \sqrt{2\overline{v}^*})$ they would choose under *laissez faire* are harmed by regulation. In other words, with endogenous service quality, default regulation entails a trade-off between social welfare and consumer surplus.

The regulator finds it optimal to allow the platform to collect more data for free when default regulation induces a quality under-provision problem. Finally, the welfare-optimal hard-cap regulation, aimed at protecting naive consumers from data exploitation by capping the amount of data the platform can collect from any consumer at $\min D_1, \widehat{D}$, remains characterized as in Proposition 6 if $D_1 > D_0$, because the optimal $\overline{v}^{\mathrm{R}}$ is set as in Proposition 7 irrespective of the hard cap.[30]

6.2 Informed Consent

Absent regulation, the platform offers naive consumers a data-exploitative option and therefore benefits from serving a larger share of naive consumers if $\overline{D}$ is sufficiently large. The opposite holds if the regulator introduces a hard cap on data trade at a level close to the free option designed for sophisticated consumers: under a hard cap, naive consumers yield the same profit as the lowest-type sophisticated consumer, while high-θ sophisticated consumers are more profitable. Hence, when D_1 is close to $\sqrt{2v}$ (absent default regulation) or to D_0 (with default regulation), the platform gains more in expectation from sophisticated consumers.[31] These comparative-statics results imply that if (similar to Gabaix and Laibson, 2006) the share of sophisticated consumers depends on the platform's choice architecture (e.g., the clarity of privacy policies), a strict hard cap provides the platform with incentives to be more transparent, improving consumer awareness and welfare.

Alternatively, suppose that the regulator can enforce an informed-consent procedure (as envisioned by the GDPR), ensuring that all consumers who opt out of the default option are made aware of the privacy terms (and hence behave like sophisticated consumers).

Since consumers can still access the platform and use the default version without understanding its privacy policy, this requirement is inconsequential if the platform can set the option $(\overline{D}, T^*(\underline{\theta}))$ as the default. This is because naive consumers, recognizing that no alternative option in the menu is more attractive from their perspective (since all others entail higher transfers), will stick to this default option. When $T^*(\underline{\theta}) < 0$, the regulator can try to prevent this exploitative outcome by requiring the default option to be offered for free or, equivalently, by requiring informed consent for any option involving a nonzero transfer.

Proposition 9 (**Optimal Menu under Informed Consent Regulation**). *Suppose that the regulator requires informed consent for any option involving a nonzero transfer. If $\overline{D}$ is sufficiently large, there exists a threshold $\tilde{\gamma} < \hat{\gamma}$ such that:*

- *For $\gamma \leq \tilde{\gamma}$, the platform offers the menu described in Proposition 1 (with $(\sqrt{2v}, 0)$ as the default option), and all consumers become sophisticated;*

[30] The "full service quality" requirement is not needed for the hard cap, because (i) naive consumers are unresponsive to v and would always select D_1, and (ii) using quality degradation to steer sophisticated types $\theta \leq \theta_0(D_1)$ away from D_1 and toward a more private option would further reduce the platform profit. If instead welfare is maximized for $D_1 = D_0$, such option is chosen taking into consideration its effect on service quality, similar to Proposition 8.

[31] This result is *a fortiori* true if $D_1 = D_0$ because, when a sophisticated consumer θ self-selects into a different option $(D^{\mathrm{R}}(\theta), T^{\mathrm{R}}(\theta))$, both the consumer and the platform are better off relative to their choice of $(D_0, 0)$.

- *For $\gamma > \tilde{\gamma}$, the platform offers $(\overline{D}, 0)$ as the default option, which is selected by naive consumers, and the options $(\sqrt{2v}, 0)$ to sophisticated consumers with $\theta \leq \theta_2$ and $(D^*(\theta), T^*(\theta))$, as defined in Proposition 1, to sophisticated consumers with $\theta > \theta_2$.*

Suppose that $\overline{D}$ is large enough that the platform gains more from a naive consumer selecting $(\overline{D}, 0)$ than, on average, from a sophisticated consumer, even when the menu designed for sophisticated consumers is undistorted by the presence of naive consumers. The considered regulatory constraints imply that the platform can extract $\overline{D}$ from naive consumers if and only if (i) it sets $(\overline{D}, 0)$ as the default option, and (ii) all other options feature non-negative transfers. If $\overline{D}$ were not the default, naive consumers would never select it, because moving away from the default requires understanding the privacy terms, implying that they would act as sophisticated consumers and therefore not choose $\overline{D}$. Then, $\overline{D}$ being the default option, it must be offered for free by regulation. Moreover, if there were any option with $T < 0$, naive consumers would move away from the default. Hence, the platform can offer $(\overline{D}, 0)$ to naive consumers only if it refrains from offering negative-price options to sophisticated ones.

The constrained-optimal menu that extracts all data from naive consumers therefore involves bunching all sophisticated consumers with $\theta \leq \theta_2$ at $(\sqrt{2v}, 0)$. When γ is small, this entails losing considerable profits from trading more data with low (sophisticated) types $\theta \leq \theta_1$. The platform then offers the same menu as in the case without naive consumers instead, displaying $(\sqrt{2v}, 0)$ as the default option. Because $T^*(\underline{\theta}) < 0$, naive consumers are attracted by the negative-price option, move away from the default toward $(D^*(\underline{\theta}), T^*(\underline{\theta}))$, understand the privacy terms, and end up behaving as sophisticated consumers.

In these circumstances, the enforcement of informed consent regulation fully addresses the issue of data exploitation and mitigates the privacy over-provision for low-θ sophisticated consumers, improving upon the *laissez-faire* outcome. Yet, since the menu is still inefficiently distorted for rent-extraction purposes, privacy-enhancing default regulation might further improve welfare.[32]

Conversely, for larger values of γ, the solution under the informed-consent constraint is weakly less efficient (and worse for consumers) than *laissez faire*: naive consumers reveal maximal data, but now receive no compensation, and the gains from trading data $D > \sqrt{2v}$ with low-θ sophisticated consumers are not realized.[33]

However, this outcome can be improved upon by default regulation. Mandating the platform to set $D = \sqrt{2v}$ as the default option leads to the *laissez-faire* outcome with all sophisticated consumers (Proposition 2), since naive consumers cannot be steered away from this free default without becoming aware of the privacy terms. Then, as above, mandating $D_0 < \sqrt{2v}$ can further improve welfare.

[32] Note that regulation must specify that $D_0 < \sqrt{2v}$ be offered as a free default option. Otherwise, as the platform cannot extract $\sqrt{2v}$ for free anymore, it might set $(\overline{D}, 0)$ as the default, because the reduced average profitability from sophisticated consumers could make it more profitable exploiting naive consumers.

[33] This result aligns with empirical findings from GDPR implementation, where privacy-conscious consumers' decision to opt out increased the traceability and advertising value of remaining consumers (Aridor et al., 2023), demonstrating how regulation-enabled privacy choices create externalities that affect the composition and value of firms' consumer base.

Proposition 10 (**Optimal Informed Consent and Default Regulation**). *Suppose that the regulator requires informed consent for consumers who move away from the default option, even if no monetary transfers are involved. Then, under the same assumptions of Proposition 5, the optimal regulatory default is* D_0^W*—the welfare-optimal default option when all consumers are sophisticated—defined in equation* (17) *for the case of* $\gamma = 0$.

Irrespective of the share of naive consumers, once the regulator requires informed consent for any option that departs from the default, movers internalize privacy losses and cannot be exploited. The platform's problem is therefore identical to the case in which all consumers are sophisticated, and it implements the menu $(D^{\mathrm{R}}(\theta))$ in (11) evaluated at $(\gamma = 0)$. Since the default is the least profitable option, the platform finds it optimal to post negative prices on data-rich alternatives to draw users away from the default, even though informed consent makes those users aware of privacy losses. With consumer naiveté no longer exploitable, the platform has no incentive to distort the menu as under *laissez faire*. Consequently, the welfare-optimal default regulation coincides with Proposition 6 evaluated at $(\gamma = 0)$; moreover, because any collection beyond the default requires informed consent, a hard cap on data is unnecessary.

7 Conclusion

We have developed a model showing how privacy regulation shapes a platform's incentives to collect user data when consumers differ in their valuations of privacy and in their understanding of data practices. We have shown that when all consumers are sophisticated, a platform may optimally *buy data* from privacy-insensitive users and *sell privacy* to sensitive ones, creating an endogenous market for personal information. However, when a large share of users fail to internalize privacy losses, the platform finds it optimal to eliminate that market altogether. By steering naive users to a data-intensive free option and avoiding any negative-price contracts, the platform avoids the payments it would otherwise owe to naive consumers who are unaware of what they are selling.

The regulatory interventions implemented or proposed in the EU data protection law that constrain the platform's design of this menu can both protect naive users and restore efficiency in the trade of data, but only within limits. A free, privacy-protecting default and a hard cap on data collection jointly discipline the platform's use of personal information: the default raises the privacy floor and induces data payments, while the cap prevents exploitation. These tools can stimulate the creation of a market in which consumers are compensated for data sharing. Yet, when the share of naive consumers is large, the level of privacy required to sustain such a market becomes too inefficient to enforce. Beyond this threshold, the regulator optimally sets the default equal to the cap and forgoes data trade altogether, accepting a uniform, privacy-preserving outcome.

Finally, we have shown that additional policy instruments—uniform service quality and informed consent regulation—complement effective privacy policy. Quality regulation prevents platforms from degrading the value of the free option to undermine the impact of default regulation. As for effective informed-consent requirements, if the platform clearly explains what sharing personal data means so that it can obtain genuine consent from users, this process turns naive users into sophisticated ones. This causes a profound change in platform strategies for buying data from users because naive users are harder to exploit. However, neither of these instruments can substitute for a combination of free defaults and data caps, because only these two policies directly determine the scope of data collection and ensure the existence of a functioning data market.

The policy implications of the model are substantial. Both *laissez-faire* and regulated versions of digital services that vary by price and privacy can create markets in an important resource of the new economy: personal data. As the importance of data continues to grow, designing these markets to maximize social welfare is an important goal. In addition, allowing as many consumers as possible to benefit from choices concerning their personal data and protecting those liable to be exploited requires careful study. The model presented here contributes to understanding the impact of such policy choices by analyzing settings very close to the requirements of the EU's Digital Markets Act, Digital Services Act, and GDPR.

We believe that studying markets for personal data is worthwhile despite the fact that subsidies to consumers are rare at present. The aggregate surplus at stake is extremely large—on the order of magnitude of the profits earned by digital platforms. These large platforms have invested in designing choice architecture and testing infrastructure as part of their ordinary course of business. Moreover, the technology is already in place—and growing with AI and distributed ledgers—to enable secure and verifiable small transactions, so that consumers could conceivably be paid for their data. For all these reasons, we expect personal data markets to develop in future. Economic research on optimal regulation in this area is therefore both intellectually challenging and valuable to policy makers.

Appendix

A Proofs

Proofs of Proposition 1. For each consumer with type θ, the platform maximizes $\alpha D + T$ subject to the individual-rationality constraint $\theta[v - L(D)] - T \geq 0$. Hence, the platform optimally sets $T = \theta[v - L(D)]$, which, substituted back into the platform profit, implies that the platform chooses the data option that maximizes $W(\theta)$ in (1). Solving this maximization problem gives $D^{\text{CI}}(\theta)$. □

Proof of Proposition 2. Statements 1–4 follow from the proofs of the corresponding parts of Proposition 3 by setting $\gamma = 0$. □

Proof of Proposition 3. We have argued in the text that the platform optimally offers the option $(\overline{D}, T(\underline{\theta}))$ to naive consumers, which establishes the result in part 5.

Proof of 1. Substituting consumers' rent obtained from (IC), with $U(\theta_1) = U(\theta_2) = 0$, and integrating by parts, we can rewrite the screening problem as

$$\max_{\{D(\theta)\}_{\theta\in\Theta},\theta_1,\theta_2} (1-\gamma)\Bigg\{\int_{\underline{\theta}}^{\theta_1}\left[\alpha D(\theta) + \left(\theta + \frac{F(\theta)}{f(\theta)}\right)\left(v - \frac{1}{2}D(\theta)^2\right)\right]\mathrm{d}F(\theta) + \int_{\theta_1}^{\theta_2}\alpha\sqrt{2v}\,\mathrm{d}F(\theta)$$

$$+\int_{\theta_2}^{\overline{\theta}}\left[\alpha D(\theta) + \left(\theta - \frac{1-F(\theta)}{f(\theta)}\right)\left(v - \frac{1}{2}D(\theta)^2\right)\right]\mathrm{d}F(\theta)\Bigg\} + \gamma\left(\alpha\overline{D} + T(\underline{\theta})\right), \tag{21}$$

where, substituting $U(\underline{\theta})$ using (IC), $T(\underline{\theta})$ writes as

$$T(\underline{\theta}) = \underline{\theta}\left(v - \frac{1}{2}D(\underline{\theta})^2\right) + \int_{\underline{\theta}}^{\theta_1}\left(v - \frac{1}{2}D(\theta)^2\right)\mathrm{d}\theta.$$

This optimization problem is only subject to (M), as (IC) has been already substituted into the objective to compute the rents, and (IR) always holds because $U(\theta) = 0$ for all $\theta \in [\theta_1, \theta_2]$, and $U(\theta) > 0$ for all other types. Neglecting the monotonicity constraint, we can maximize pointwise w.r.t. $D(\theta)$. For all $\gamma > 0$, the derivative w.r.t. D has a downward discontinuity at $\theta = \underline{\theta}$, which, as argued in the text, implies that the solution must exhibit bunching for $\theta \in [\underline{\theta}, \theta_0]$, with $\theta_0 \in (\underline{\theta}, \theta_1)$, which we shall check ex-post. Denote by $(\widehat{D}, \widehat{T})$ the option offered in this interval. Using (IC) and the fact that $U(\theta_1) = 0$, integration by parts yields

$$\int_{\theta_0}^{\theta_1} U(\theta)\,\mathrm{d}F(\theta) = -\int_{\theta_0}^{\theta_1}\frac{F(\theta) - F(\theta_0)}{f(\theta)}\left(v - \frac{1}{2}D(\theta)^2\right)\mathrm{d}F(\theta),$$

and so the platform's objective rewrites as

$$(1-\gamma)\left\{\int_{\underline{\theta}}^{\theta_0}\left(\alpha\widehat{D}+\widehat{T}\right)\mathrm{d}F(\theta)+\int_{\theta_0}^{\theta_1}\left[\alpha D(\theta)+\left(\theta+\frac{F(\theta)-F(\theta_0)}{f(\theta)}\right)\left(v-\frac{1}{2}D(\theta)^2\right)\right]\mathrm{d}F(\theta)\right.$$

$$\left.+\int_{\theta_1}^{\theta_2}\alpha\sqrt{2v}\,\mathrm{d}F(\theta)+\int_{\theta_2}^{\overline{\theta}}\left[\alpha D(\theta)+\left(\theta-\frac{1-F(\theta)}{f(\theta)}\right)\left(v-\frac{1}{2}D(\theta)^2\right)\right]\mathrm{d}F(\theta)\right\}+\gamma\left(\alpha\overline{D}+\widehat{T}\right),$$

where, since $\widehat{T}$ is selected by type $\theta=\theta_0$, it can be written using (IC) as

$$\widehat{T}=\theta_0\left(v-\frac{1}{2}\widehat{D}^2\right)+\int_{\theta_0}^{\theta_1}\left(v-\frac{1}{2}D(\theta)^2\right)\mathrm{d}\theta. \tag{22}$$

Maximizing pointwise w.r.t. $D(\theta)$ in the two intervals (θ_0,θ_1) and $(\theta_2,\overline{\theta}]$ yields the schedule $D^*(\theta)$ for these intervals given in (6) which, under the considered monotone virtual valuations assumptions, satisfies the omitted constraint (M).

Maximizing w.r.t. θ_1 and θ_2 yields that the menu must be continuous at $\theta=\theta_1$ and $\theta=\theta_2$, where $D^*(\theta)=\sqrt{2v}$, which gives (8). The monotonicity of $D^*(\cdot)$ and these continuity conditions at $\theta\in\{\theta_1,\theta_2\}$ imply that θ_1 and θ_2 are unique. Similarly, maximizing w.r.t. θ_0 yields the continuity condition $\lim_{\theta\to\theta_0}D^*(\theta)=\widehat{D}$, so that θ_0 is the function of $\widehat{D}$ defined as

$$\frac{\alpha}{\theta_0+\frac{F(\theta_0)}{f(\theta_0)}+\frac{\gamma}{(1-\gamma)f(\theta_0)}}=\widehat{D}. \tag{23}$$

The FOC of the platform problem w.r.t. $\widehat{D}$ simplifies as

$$(1-\gamma)F(\theta_0)(\alpha-\theta_0\widehat{D})-\gamma\theta_0\widehat{D}=0.$$

Substituting $\widehat{D}$ from (23) and simplifying yields (7).

Proof of 2. From $T(\theta)=\theta(v-[D(\theta)]^2/2)-U(\theta)$ we have $\dot{T}(\theta)=-\theta D(\theta)\dot{D}(\theta)\geq 0$ because $\dot{D}(\theta)\geq 0$ from (M), with strict inequality whenever the allocation is separating, and $T^*(\theta)=0$ for $\theta\in[\theta_1,\theta_2]$ given that for these types $D^*(\theta)=\sqrt{2v}$ and $U^*(\theta)=0$, which yields the properties of $T^*(\theta)$ given in the statement.

Proof of 3 and 4. We have shown above that the cutoffs θ_0, θ_1, and θ_2 are pinned down by the continuity of the optimal menu, yielding (7) and (8).

For $\gamma=0$, as $F(\theta)/f(\theta)$ is increasing, $\theta_1>\underline{\theta}$ if and only if $\underline{\theta}<\alpha/\sqrt{2v}$. Similarly, irrespective of γ, as $[1-F(\theta)]/f(\theta)$ is decreasing, $\theta_2<\overline{\theta}$ if and only if $\overline{\theta}>\alpha/\sqrt{2v}$. Moreover, $\theta+F(\theta)/f(\theta)>\theta-[1-F(\theta)]/f(\theta)$ implies that $\theta_1<\theta_2$ for $\gamma=0$; this is *a fortiori* true for $\gamma>0$ given that θ_2 is not a function of γ whereas θ_1 is decreasing in γ because. To see this, note that for $\theta\in(\theta_0,\theta_1)$, we have that $D^*(\theta)$ is decreasing in γ:

$$\operatorname{sgn}\left[\frac{\partial D^*(\theta|\theta\in(\theta_0,\theta_1))}{\partial\gamma}\right]=\operatorname{sgn}\left[-\frac{1}{f(\theta)}\right]<0,$$

which, together with the continuity of the optimal mechanism—specifically, $\lim_{\theta \to \theta_1} D^*(\theta) = D^*(\theta_1) = \sqrt{2v}$, independent of γ—implies that θ_1 is a decreasing function of γ.

However, the considered solution is admissible if and only if $\widehat{D} > \sqrt{2v}$. Otherwise, the optimal menu—being decreasing and continuous—exhibits bunching at $D^*(\theta) = \sqrt{2v}$ for all $\theta \in [\underline{\theta}, \theta_2]$ (in these circumstances, we denote $\underline{\theta} = \theta_0 = \theta_1$).

To show that the condition $\widehat{D} > \sqrt{2v}$ is equivalent to $\gamma < \hat{\gamma}$, let us first show that the data option $\widehat{D}$ is decreasing in γ. To this end, note that we can rewrite the platform's expected profit as $\Pi := \pi_S(\widehat{D}) + \gamma[\pi_N(\widehat{D}) - \pi_S(\widehat{D})]$, where

$$\pi_S(\widehat{D}) := \int_{\underline{\theta}}^{\theta_0(\widehat{D})} [\alpha\widehat{D} + \widehat{T}(\widehat{D})] \, \mathrm{d}F(\theta) + \int_{\theta_0(\widehat{D})}^{\overline{\theta}} [\alpha D^*(\theta) + T^*(\theta)] \, \mathrm{d}F(\theta) \quad \text{and} \quad \pi_N(\widehat{D}) := \alpha\overline{D} + \widehat{T}(\widehat{D})$$

are the profits that the platform makes from each sophisticated and naive consumer, respectively, with $\widehat{T}(\widehat{D})$ and $\theta_0(\widehat{D})$ being the functions defined in (22) and (23). We then have $\pi'_S = F(\theta_0)[\alpha - \theta_0\widehat{D}] > 0$, because, using (23), $\alpha - \theta_0\widehat{D} = \widehat{D}[F(\theta_0)/f(\theta_0) + \gamma/[(1-\gamma)f(\theta_0)]] > 0$, and $\pi'_N = \widehat{T}' = -\theta_0\widehat{D} < 0$. Therefore, the platform profit Π has decreasing differences in γ and $\widehat{D}$, and a standard monotone comparative statics argument (Milgrom and Shannon, 1994) establishes that $\widehat{D}$ is decreasing in γ.

Then, since (i) $\widehat{D}$ is decreasing in γ, (ii) for $\gamma \to 0$, $\theta_0 \to \underline{\theta}$ and so $\widehat{D} \to \alpha/\underline{\theta} > \sqrt{2v}$ if and only if $\underline{\theta} < \alpha/\sqrt{2v}$, and (iii) for $\gamma \to 1$, $\widehat{D} \to 0 < \sqrt{2v}$, it follows that the condition $\widehat{D} > \sqrt{2v}$ is satisfied if and only if $\gamma < \hat{\gamma}$, where $\hat{\gamma} \in [0, 1)$ is a decreasing function of v and such that $\hat{\gamma} = 0$ if and only if $\underline{\theta} \geq \alpha/\sqrt{2v}$. □

Proof of Corollary 1. The results in Proposition 3 imply that all data options $D \leq \sqrt{2v}$ do not depend on γ. Since θ_1 is decreasing in γ, this implies that $D^*(\theta|\theta \geq \theta_1|_{\gamma=0})$ does not depend on γ. In what follows, we consider $\theta < \theta_1|_{\gamma=0}$ and $\gamma < \hat{\gamma}$.

We have shown in the proof of Proposition 3 that $D^*(\theta|\theta \leq \theta_1)$ decrease in γ. As for consumer rents and transfers, from (IC) we have that $U^*(\theta|\theta < \theta_1) = -\int_{\theta}^{\theta_1} [v - [D^*(\tilde{\theta})]^2/2] \, \mathrm{d}\tilde{\theta}$ is increasing in $D^*(\tilde{\theta})$, and so $T^*(\theta) = \theta[v - [D^*(\theta)]^2/2] - U^*(\theta)$ decreases with $D(\tilde{\theta})$, for all $\tilde{\theta} \geq \theta$. As these data options are decreasing in γ, we have that $U^*(\theta)$ is decreasing and $T^*(\theta) < 0$ is increasing—hence $|T^*(\theta)|$ is decreasing—in γ. Finally, since naive consumers select $(\overline{D}, T^*(\underline{\theta}))$, this result also implies that their rent is decreasing in γ. □

Proof of Proposition 4. Since the platform offers $(\overline{D}, T^*(\underline{\theta}))$ to naive consumers in the absence of regulation, any hard cap $D_1 < \overline{D}$ binds for naive consumers, implying that $D = D_1$ will be offered to these consumers. Recalling that $D^*(\theta) = \widehat{D}$ for all $\theta \in [\underline{\theta}, \theta_0]$, we have:

- If $D_1 \geq \widehat{D}$, the optimal menu designed for sophisticated consumers, which features a decreasing data schedule, satisfies the hard-cap constraint $D^*(\theta) \leq D_1$ for all θ. Then, since optimization w.r.t. $D(\theta)$ does not depend on the data option offered to naive consumers, hard-cap regulation does not affect the menu offered to sophisticated consumers, and naive consumers obtain $(D_1, T^*(\underline{\theta}))$.

- If $\sqrt{2v} \le D_1 < \widehat{D}$, the decreasing data schedule $D^*(\theta)$ does not satisfy the hard-cap constraint for all $\theta \le \theta_0(D_1)$, where, the menu being continuous, the cutoff $\theta_0(D_1) > \theta_0$ is obtained from

$$\frac{\alpha}{\theta_0 + \frac{F(\theta_0)}{f(\theta_0)} + \frac{\gamma}{(1-\gamma)f(\theta_0)}} = D_1,$$

  and is thus decreasing in D_1 and such that $\lim_{D_1 \to \widehat{D}} \theta_0(D_1) = \theta_0$ defined in (7) and $\theta_0(\sqrt{2v}) = \theta_1$ defined in Eq. (8). Since the objective in (21) is concave in each $D(\theta)$, further distorting the menu would decrease the platform profit. Hence, the platform offers to sophisticated consumers the data options $D(\theta) = \min\{D_1, D^*(\theta)\}$ and the access prices $T(\theta) = \max\{T^*(\theta_0(D_1)), T^*(\theta)\}$, where $T^*(\theta_0(D_1)) \in (T^*(\theta|\theta < \theta_0(D_1)), 0)$. Finally, the option $(D_1, T^*(\theta_0(D_1)))$ is offered also to naive consumers.

- If $D_1 < \sqrt{2v}$, the decreasing data schedule $D^*(\theta)$ does not satisfy the hard-cap constraint for all $\theta \in [\underline{\theta}, \theta_2(D_1)]$, where for all $D_1 \in (\alpha/\overline{\theta}, \sqrt{2v})$ the cutoff $\theta_2(D_1)$ is obtained from the continuity condition $D^*(\theta_2) = D_1$, hence is decreasing in D_1 and such that $\lim_{D_1 \to \sqrt{2v}} \theta_2(D_1) = \theta_2$ defined in (8), whereas $\theta_2(D_1) = \overline{\theta}$ for all $D_1 \le \alpha/\overline{\theta}$. As above, because the objective in (21) is concave in each $D(\theta)$, the platform sets $D(\theta) = \min\{D_1, D^*(\theta)\}$. Yet, here there are two possibilities to consider. If the platform offers $(D_1, 0)$ to naive consumers and sophisticated consumers with type $\theta \le \theta_2(D_1)$ and $(D^*(\theta), T^*(\theta))$ to sophisticated types $\theta > \theta_2(D_1)$, it makes profit

$$\Pi_f(D_1) := \gamma\alpha D_1 + (1-\gamma)\int_{\underline{\theta}}^{\overline{\theta}} [\alpha D(\theta) + T(\theta)]\,\mathrm{d}F(\theta),$$

  where $D(\theta) = \min\{D_1, D^*(\theta)\}$, $T(\theta|\theta \le \theta_2(D_1)) = 0$ and $T(\theta|\theta > \theta_2(D_1)) = \theta(v - \frac{1}{2}[D^*(\theta)]^2) - \theta_2(D_1)(v - \frac{1}{2}D_1^2) - \int_{\theta_2(D_1)}^{\theta}(v - \frac{1}{2}[D^*(\tilde{\theta})]^2)\,\mathrm{d}\tilde{\theta}$. Alternatively, the platform can offer D_1 at a positive price $T_1 > 0$, which implies excluding naive consumers and sophisticated consumers $\theta \le \widehat{\theta}(T_1)$, where $\widehat{\theta}(T_1) := T_1/(v - D_1^2/2)$. In this case, by optimally choosing T_1, the platform makes profit

$$\Pi_s := \max_{T_1 > 0}\,(1-\gamma)\int_{\widehat{\theta}(T_1)}^{\overline{\theta}} [\alpha D(\theta) + T(\theta) + T_1]\,\mathrm{d}F(\theta),$$

  because the platform optimally increases by the same T_1 the payment to each participating type. Both profit functions are increasing in D_1, and

$$\frac{\partial \Pi_f}{\partial D_1} = \gamma\alpha + (1-\gamma)\int_{\underline{\theta}}^{\overline{\theta}} \frac{\partial}{\partial D_1}[\alpha D(\theta) + T(\theta)]\,\mathrm{d}F(\theta)$$

is strictly larger than

$$\frac{\partial \Pi_s}{\partial D_1} = (1-\gamma)\left[\int_{\widehat{\theta}}^{\overline{\theta}} \frac{\partial}{\partial D_1}[\alpha D(\theta)+T(\theta)]\,\mathrm{d}F(\theta) - \frac{\partial \widehat{\theta}}{\partial D_1} f(\widehat{\theta})(\alpha D_1 + T_1)\right],$$

because the integrand function and $\partial\widehat{\theta}/\partial D_1$ are both positive. At $D_1 \to \sqrt{2v}$, we have $\Pi_f(D_1) > \Pi_s(D_1)$ because $v - D_1^2/2 \to 0$ and so $T_1 \to 0$, which implies that it is never optimal to exclude naive consumers. Conversely, at $D_1 = 0$, $\Pi_f(0) = 0 < \max_{T_1} T_1[1 - F(\widehat{\theta}(T_1))] = \Pi_s(0)$. Therefore, we can conclude that there exists $\widehat{D}_1 \in (0, \sqrt{2v})$ such that the platform offers D_1 for free if and only if $D_1 \geq \widehat{D}_1$, and switches to a pure subscription model otherwise. By the implicit function theorem,

$$\frac{\partial \widehat{D}_1}{\partial \gamma} = -\frac{\frac{\partial \Pi_f}{\partial \gamma} - \frac{\partial \Pi_s}{\partial \gamma}}{\frac{\partial \Pi_f}{\partial D_1} - \frac{\partial \Pi_s}{\partial D_1}},$$

whose denominator is positive, as shown above. Denoting by T_1^* the profit-maximizing price in Π_s, since data policies and transfers in both solutions do not depend on γ, we have that the numerator equals

$$\left(\alpha D_1 - \int_{\underline{\theta}}^{\overline{\theta}} [\alpha D(\theta) + T(\theta)]\,\mathrm{d}F(\theta)\right) + \int_{\widehat{\theta}(T_1^*)}^{\overline{\theta}} [\alpha D(\theta) + T(\theta) + T_1^*]\,\mathrm{d}F(\theta) \geq \alpha D_1 > 0,$$

where the first inequality follows from Π_s being higher at T_1^* than at $T_1 = 0$ (for which $\widehat{\theta} = \underline{\theta}$). Therefore, we can conclude that $\widehat{D}_1$ decreases with γ.

Summing up, for all $D_1 < \overline{D}$, the platform offers $D(\theta) = \min\{D_1, D^*(\theta)\}$ to sophisticated consumers (though some low types might be excluded if D_1 is sufficiently small). □

Proof of Proposition 5. We first argue that incentive compatibility requires all consumers to receive rents at least as large as under the regulated default option, i.e., $U(\theta) \geq U_0(\theta)$ for all θ. Indeed, if the platform were to offer a menu such that $U(\tilde{\theta}) < U_0(\tilde{\theta})$ for some type $\tilde{\theta}$, that type would instead prefer the regulated default option $(D_0, 0)$ to the menu option $(D(\tilde{\theta}), T(\tilde{\theta}))$ intended for them. Then, the monotonicity condition (M) implies that the regulatory option can be taken by a (possibly non-degenerate) interval of types $\theta \in [\theta_1, \theta_2]$, with $\underline{\theta} \leq \theta_1$ and $\theta_2 \leq \overline{\theta}$. For all $\theta < \theta_1$, $D(\theta) > D_0$ and so differentiating (IC) we have

$$\dot{U}(\theta) = v - \frac{1}{2}[D(\theta)]^2 < \dot{U}_0(\theta) = v - \frac{1}{2}D_0^2,$$

which implies that $U(\theta) > U_0(\theta)$. Similarly, for all $\theta > \theta_2$, $D(\theta) < D_0$ and differentiating (IC) gives

$$\dot{U}(\theta) = v - \frac{1}{2}[D(\theta)]^2 > \dot{U}_0(\theta) = v - \frac{1}{2}D_0^2,$$

which implies that again $U(\theta) > U_0(\theta)$. Therefore, $U(\theta) = U_0(\theta)$ for $\theta \in [\theta_1, \theta_2]$ and $U(\theta) > U_0(\theta)$ otherwise, with

$$U(\theta) = U_0(\theta_1) - \int_{\theta}^{\theta_1} \left(v - \frac{1}{2}D(\tilde{\theta})^2\right) \mathrm{d}\tilde{\theta} \quad \forall \theta \in [\underline{\theta}, \theta_1),$$

and

$$U(\theta) = U_0(\theta_2) + \int_{\theta_2}^{\theta} \left(v - \frac{1}{2}D(\tilde{\theta})^2\right) \mathrm{d}\tilde{\theta} \quad \forall \theta \in (\theta_2, \overline{\theta}].$$

Moreover, as argued in the text, the platform still offers $(\overline{D}, T(\underline{\theta}))$ to naive consumers (which establishes the result in part 5).

Proof of 1. Substituting the rents derived above into the platform's objective and integrating by parts, the regulated platform's problem writes as

$$\max_{\{D(\theta)\}_{\theta\in\Theta}, \theta_1, \theta_2} (1-\gamma)\left\{ \int_{\underline{\theta}}^{\theta_1} \left[\alpha D(\theta) + \left(\theta + \frac{F(\theta)}{f(\theta)}\right)\left(v - \frac{1}{2}D(\theta)^2\right) - U_0(\theta_1)\right] \mathrm{d}F(\theta) + \int_{\theta_1}^{\theta_2} \alpha D_0 \, \mathrm{d}F(\theta) \right.$$

$$\left. + \int_{\theta_2}^{\overline{\theta}} \left[\alpha D(\theta) + \left(\theta - \frac{1-F(\theta)}{f(\theta)}\right)\left(v - \frac{1}{2}D(\theta)^2\right) - U_0(\theta_2)\right] \mathrm{d}F(\theta)\right\} + \gamma\left(\alpha \overline{D} + T(\underline{\theta})\right),$$

where $T(\underline{\theta}) = \underline{\theta}\left(v - \frac{1}{2}D(\underline{\theta})^2\right) - U(\underline{\theta})$. This problem is subject only to (M), as (IC) has been already substituted into the objective to compute the rents, and (IR) always holds given that $U(\theta) \geq U_0(\theta) > 0$ for all θ. Neglecting the monotonicity constraint, we can maximize pointwise w.r.t. $D(\theta)$ in the two intervals $[\underline{\theta}, \theta_1)$ and $(\theta_2, \overline{\theta}]$. As under *laissez faire*, for all $\gamma > 0$ this solution would violate (M) at $\theta = \underline{\theta}$, and so (M) must bind in an interval $\theta \in [\underline{\theta}, \theta_0]$, where the platform offers a bunching option $(\widehat{D}, \widehat{T})$. Using the fact that this bunching option is selected by type θ_0,

$$\widehat{T} = \theta_0\left(v - \frac{1}{2}\widehat{D}^2\right) - U_0(\theta_1) + \int_{\theta_0}^{\theta_1}\left(v - \frac{1}{2}D(\theta)^2\right) \mathrm{d}\theta. \tag{24}$$

Then, by the same steps as in the proof of Proposition 3, the platform's objective becomes

$$(1-\gamma)\left\{ \int_{\underline{\theta}}^{\theta_0} \left[\alpha\widehat{D} + \widehat{T}\right] \mathrm{d}F(\theta) + \int_{\theta_0}^{\theta_1} \left[\alpha D(\theta) + \left(\theta + \frac{F(\theta) - F(\theta_0)}{f(\theta)}\right)\left(v - \frac{1}{2}D(\theta)^2\right) - U_0(\theta_1)\right] \mathrm{d}F(\theta) \right.$$

$$\left. + \int_{\theta_1}^{\theta_2} \alpha D_0 \, \mathrm{d}F(\theta) + \int_{\theta_2}^{\overline{\theta}} \left[\alpha D(\theta) + \left(\theta - \frac{1-F(\theta)}{f(\theta)}\right)\left(v - \frac{1}{2}D(\theta)^2\right) - U_0(\theta_2)\right] \mathrm{d}F(\theta)\right\} + \gamma\left(\alpha\overline{D} + \widehat{T}\right).$$

Optimizing pointwise with respect to the data options we find the schedule $D^{\mathrm{R}}(\theta)$ given in (11) which, under the considered monotone virtual valuations assumptions, satisfies the omitted constraint (M). Maximizing w.r.t. θ_0, θ_1, and θ_2 yields that the menu must be continuous: θ_0 is as under *laissez faire* (i.e., solves (7)), whereas $\theta_1(D_0)$ and $\theta_2(D_0)$ satisfy (12).

Proof of 2. Since $\dot{T}^{\mathrm{R}}(\theta) = -\theta D^{\mathrm{R}}(\theta)\dot{D}^{\mathrm{R}}(\theta) \geq 0$, with strict inequality (hence, $T^{\mathrm{R}}(\theta)$ strictly increasing) if and only if the allocation is separating, we have that $T^{\mathrm{R}}(\theta) = \dot{T}^{\mathrm{R}}(\theta) = 0$ for $\theta \in [\theta_1(D_0), \theta_2(D_0)]$, while $T^{\mathrm{R}}(\theta)$ is strictly negative (resp. positive) for $\theta < \theta_1(D_0)$ (resp. $\theta > \theta_2(D_0)$).

Proof of 3–4. We have shown above that the cutoffs θ_0, $\theta_1(D_0)$, and $\theta_2(D_0)$ are pinned down by the continuity of the optimal menu. Monotonicity and continuity of the menu imply that both $\theta_1(D_0)$ and $\theta_2(D_0)$ are unique and decreasing in D_0, with $\theta_1(D_0) < \theta_2(D_0)$, and $\theta_2(D_0) < \overline{\theta}$ for all $D_0 < \alpha/\overline{\theta}$.

As in Proposition 3, the solution characterized above is admissible if and only if $\widehat{D} > D_0$. Since $\widehat{D}$ does not depend on D_0, and, as shown in the proof of Proposition 3 (i) it is decreasing in γ, (ii) $\widehat{D} \to \alpha/\underline{\theta}$ as $\gamma \to 0$, and (iii) $\widehat{D} \to 0$ as $\gamma \to 1$, it follows that the condition $\widehat{D} > D_0$ is satisfied if and only if $\gamma < \hat{\gamma}(D_0)$, where, provided that $\underline{\theta} < \alpha/\sqrt{2v}$, $\hat{\gamma}(D_0) \in (0, 1)$ is such that $\widehat{D} = D_0$; if instead $\underline{\theta} > \alpha/\sqrt{2v}$, $\hat{\gamma}(D_0) = 0$ for some D_0 not too low relative to $\sqrt{2v}$. $\widehat{D}$ being independent of D_0 also implies that $\hat{\gamma}(D_0)$ is decreasing in D_0. Then, for all $D_0 \in (\alpha/\overline{\theta}, \sqrt{2v})$, $\underline{\theta} < \theta_0 < \theta_1(D_0) < \theta_2(D_0) < \overline{\theta}$ for $\gamma \in [0, \hat{\gamma}(D_0))$, whereas the interval $[\underline{\theta}, \theta_1(D_0))$ is empty—hence, we denote $\underline{\theta} = \theta_0 = \theta_1(D_0)$—for $\gamma \geq \hat{\gamma}(D_0)$. □

Proof of Corollary 2. The *laissez-faire* solution in Proposition 3 coincides with the regulated platform's solution in Proposition 5 if the default option equals the free data option absent regulation: $D_0 = \sqrt{2v}$. Therefore, the cutoffs θ_1 and θ_2 defined under *laizzez faire* coincide with $\theta_1(\sqrt{2v})$ and $\theta_2(\sqrt{2v})$. As $\theta_1(D_0)$ and $\theta_2(D_0)$ are decreasing functions of D_0, for all $D_0 \in (\alpha/\overline{\theta}, \sqrt{2v})$ we have $\theta_1(\sqrt{2v}) \leq \theta_1(D_0)$, with strict inequality for all $\gamma < \hat{\gamma}(D_0)$ (otherwise, both cutoffs coincide with $\underline{\theta}$) and $\theta_2(\sqrt{2v}) < \theta_2(D_0)$.

Then, comparing the data options $D^*(\theta)$ in (6) and $D^{\mathrm{R}}(\theta)$ in (11) yields that $D^*(\theta) = D^{\mathrm{R}}(\theta)$ for $\theta \leq \theta_1(\sqrt{2v})$ and $\theta \geq \theta_2(D_0)$, and $D^*(\theta) > D^{\mathrm{R}}(\theta)$ for $\theta_1(\sqrt{2v}) < \theta < \theta_2(D_0)$ (with $\theta_1(\sqrt{2v}) = \underline{\theta}$ if $\gamma \geq \hat{\gamma}(\sqrt{2v})$). Specifically, if $\theta_1(D_0) < \theta_2(\sqrt{2v})$: (i) types $\theta_1(\sqrt{2v}) < \theta < \theta_1(D_0)$ (this region being empty for $\gamma \geq \hat{\gamma}(D_0)$) are bunched at $\sqrt{2v}$ under *laissez faire* but receive the lower separating data options $D^{\mathrm{R}}(\theta) \in (D_0, \sqrt{2v})$ under regulation, (ii) types $\theta_1(D_0) \leq \theta \leq \theta_2(\sqrt{2v})$ are bunched at the zero-price option in both solutions, which is more private under regulation, and (iii) types $\theta_2(\sqrt{2v}) < \theta < \theta_2(D_0)$ are bunched at D_0 under regulation but receive the higher separating data options $D^*(\theta) \in (D_0, \sqrt{2v})$ under *laissez faire*; if instead $\theta_1(D_0) > \theta_2(\sqrt{2v})$: (i) types $\theta_1(\sqrt{2v}) < \theta \leq \theta_2(\sqrt{2v})$ are bunched at $\sqrt{2v}$ under *laissez faire* but receive the lower separating data options $D^{\mathrm{R}}(\theta) \in (D_0, \sqrt{2v})$ under regulation, (ii) types $\theta_2(\sqrt{2v}) < \theta < \theta_1(D_0)$ receive separating allocations in both solutions, which are more private under regulation, and (iii) types $\theta_1(D_0) \leq \theta < \theta_2(D_0)$ are bunched at D_0 under regulation but receive the higher separating data option $D^*(\theta) \in (D_0, \sqrt{2v})$ under *laissez faire*.

The comparison of access prices is as follows. Suppose, first, that $\theta_1(D_0) \leq \theta_2(\sqrt{2v})$. Then, for all $\theta \in [\theta_1(D_0), \theta_2(\sqrt{2v})]$, $T^{\mathrm{R}}(\theta) = 0 = T^*(\theta)$, and:

1. For lower types, $\theta \in [\underline{\theta}, \theta_1(D_0))$ (this interval being non-empty if $\gamma < \hat{\gamma}(D_0)$):
    - for $\theta \in [\theta_1(\sqrt{2v}), \theta_1(D_0))$, $T^{\mathrm{R}}(\theta) < 0 = T^*(\theta)$;

- for $\theta \in [\underline{\theta}, \theta_1(\sqrt{2v}))$ (this interval being non-empty if $\gamma < \hat{\gamma}(\sqrt{2v})$), $D(\theta)$ and so $\dot{T}(\theta)$ is the same in the two solutions, which implies that $T^{\mathrm{R}}(\theta) < T^*(\theta)$ keeps holding.

2. For higher types, $\theta \in (\theta_2(\sqrt{2v}), \overline{\theta}]$:

- for $\theta \in (\theta_2(\sqrt{2v}), \theta_2(D_0)]$, $T^{\mathrm{R}}(\theta) = 0 < T^*(\theta)$;
- for $\theta \in (\theta_2(D_0), \overline{\theta}]$, $D(\theta)$ and so $\dot{T}(\theta)$ is the same in the two solutions, which implies that $T^{\mathrm{R}}(\theta) < T^*(\theta)$ keeps holding.

If instead $\theta_1(D_0) > \theta_2(\sqrt{2v})$, the interval of types where transfers are zero in both solutions is empty (whenever $T^{\mathrm{R}}(\theta) = 0$, $T^*(\theta) > 0$), and, by the same steps as above, $T^{\mathrm{R}}(\theta) < T^*(\theta)$ for all θ.

Finally, since all consumers receive strictly more privacy, strictly lower transfers, or both, it follows that all rents are strictly larger under regulation. □

Proof of Corollary 3. Take any D_0 such that $\hat{\gamma}(\sqrt{2v}) < \gamma < \hat{\gamma}(D_0)$. Then, the interval $[\underline{\theta}, \theta_1(D_0))$ is empty under *laissez faire* (equivalently, if $D_0 = \sqrt{2v}$) but not under regulation. Therefore, consumers $\theta \in [\underline{\theta}, \theta_1(D_0)]$ obtain $(D^*(\theta) = \sqrt{2v},\ T^*(\theta) = 0)$ under *laissez faire*, and $(D^{\mathrm{R}}(\theta) > D_0,\ T^R(\theta) < 0)$ under regulation, which establishes the *data market creation effect*.

Since $T^{\mathrm{R}}(\underline{\theta}) < 0 = T^*(\underline{\theta})$, naive consumers obtain the exploitative data option $\overline{D}$ in both scenarios, but receive a transfer from the platform only under regulation, which establishes the *compensation for data exploitation effect*. □

Proof of Proposition 6. We characterize welfare-optimal regulation assuming that the regulatory hard cap D_1 and default option D_0 are distinct and bind for sophisticated consumers, i.e., $D_0 < \sqrt{2v}$, and $D_0 < D_1 < \widehat{D}$. For all $D_0 < \sqrt{2v}$, the regulatory default option is selected by sophisticated consumers $\theta \in [\theta_1(D_0), \theta_2(D_0)]$, with the cutoffs $\theta_1(D_0)$ and $\theta_2(D_0)$ defined in (12). For all $D_1 \in (D_0, \widehat{D})$, the regulatory hard cap is selected by naive consumers and sophisticated consumers $\theta \in [\underline{\theta}, \theta_0(D_1)]$, with the cutoff $\theta_0(D_1)$ defined in (10). As a result, letting $D^{\mathrm{R}}(\theta)$ denote the optimal data schedule from Proposition 5, social welfare can be written as

$$W = (1-\gamma)\left\{ v\mathbb{E}_F[\theta] + \int_{\underline{\theta}}^{\theta_0(D_1)} \left[\alpha D_1 - \frac{\theta}{2} D_1^2\right] \mathrm{d}F(\theta) + \int_{\theta_0(D_1)}^{\theta_1(D_0)} \left[\alpha D^{\mathrm{R}}(\theta) - \frac{\theta}{2}[D^{\mathrm{R}}(\theta)]^2\right] \mathrm{d}F(\theta) \right.$$
$$\left. + \int_{\theta_1(D_0)}^{\theta_2(D_0)} \left[\alpha D_0 - \frac{\theta}{2} D_0^2\right] \mathrm{d}F(\theta) + \int_{\theta_2(D_0)}^{\overline{\theta}} \left[\alpha D^{\mathrm{R}}(\theta) - \frac{\theta}{2}[D^{\mathrm{R}}(\theta)]^2\right] \mathrm{d}F(\theta) \right\} + \gamma\left[\alpha D_1 - \mathbb{E}_{F_N}[\theta]\left(v - \frac{D_1^2}{2}\right)\right].$$

We begin by characterizing the optimal default D_0. Letting $x := \alpha / D_0$, the first order condition with respect to D_0 can be written as

$$x = \mathbb{E}[\theta \mid \theta \in [\theta_1(x), \theta_2(x)]],$$

where θ_1 and θ_2 solve

$$h_1(\theta_1) := \theta_1 + \frac{F(\theta_1)}{f(\theta_1)} + \frac{\gamma}{(1-\gamma)f(\theta_1)} = x$$

and

$$h_2(\theta_2) := \theta_1 - \frac{1 - F(\theta_2)}{f(\theta_2)} = x.$$

Because the regulator's objective is increasing in D_0 at $D_0 = \alpha/\overline{\theta}$, uniqueness of a solution to the first-order condition establishes the characterization of the optimum in (17). We begin with the following lemma.

Lemma A.1. *The map*

$$G : x \to \mathbb{E}[\theta \mid \theta \in [\theta_1(x), \theta_2(x)]]$$

is globally Lipschitz on $[\underline{\theta}, \overline{\theta}]$ *with constant*

$$L_* := \frac{M}{m} \frac{1}{2 - B}.$$

Proof of Lemma A.1. It is immediate to show that, under the regularity conditions (14)–(16) with $B < 2$, the functions h_1 and h_2 are strictly increasing on $[0, 1]$, for all γ. Moreover,

$$h_1'(\theta) \geq 2 \quad \text{and} \quad h_2'(\theta) \geq 2 - B \quad \forall\, \theta,$$

and consequently,

$$\theta_1'(x) \leq \frac{1}{2} \quad \text{and} \quad \theta_2'(x) \leq \frac{1}{2 - B}. \tag{25}$$

Now let

$$A(x) := \int_{\theta_1(x)}^{\theta_2(x)} f(\theta)\, \mathrm{d}\theta = F(\theta_2(x)) - F(\theta_1(x)) > 0,$$

so that

$$G(x) = \frac{\int_{\theta_1(x)}^{\theta_2(x)} \theta f(\theta)\, \mathrm{d}\theta}{A(x)}.$$

By the Leibniz rule,

$$G'(x) \;=\; \frac{\big(\theta_2 - G(x)\big)\, f(\theta_2)\, \theta_2'(x) + \big(G(x) - \theta_1\big)\, f(\theta_1)\, \theta_1'(x)}{A(x)}. \tag{26}$$

Because $\theta_1 \leq G(x) \leq \theta_2$, define

$$\alpha(x) := \frac{\theta_2(x) - G(x)}{\theta_2(x) - \theta_1(x)} \in [0, 1], \quad 1 - \alpha(x) = \frac{G(x) - \theta_1(x)}{\theta_2(x) - \theta_1(x)}.$$

Using (26), the bounds $f(\theta) \leq M$ and $A(x) \geq m\big(\theta_2(x) - \theta_1(x)\big)$ from (14), and (25), we obtain for every $x \in [0,1]$,

$$\begin{aligned} G'(x) &\le \frac{M}{A(x)}\Big[\big(\theta_2 - G(x)\big)\theta_2'(x) + \big(G(x) - \theta_1\big)\theta_1'(x)\Big] \\ &= \frac{M}{A(x)}\Big[\alpha(x)\big(\theta_2 - \theta_1\big)\theta_2'(x) + \big(1-\alpha(x)\big)\big(\theta_2 - \theta_1\big)\theta_1'(x)\Big] \\ &\le \frac{M}{m}\Big[\alpha(x)\,\theta_2'(x) + \big(1-\alpha(x)\big)\,\theta_1'(x)\Big] \\ &\le \frac{M}{m}\Big[\alpha(x)\,\frac{1}{2-B} + \big(1-\alpha(x)\big)\,\frac{1}{2}\Big] \le \frac{M}{m}\,\frac{1}{2-B} \;\; =: \; L_*, \end{aligned}$$

since $B < 2$ implies $(2-B)^{-1} \ge 1/2$. Thus $G'(x) \le L_*$ for all $x \in [0,1]$.

Finally, G is C^1 under the stated regularity, so by the mean value theorem,

$$|G(x) - G(y)| \;\le\; \sup_{\xi \in [0,1]} G'(\xi)\,|x-y| \;\le\; L_*\,|x-y| \qquad (\forall\, x, y \in [0,1]).$$

Thus G is globally Lipschitz on $[0,1]$ with constant $L_* = \dfrac{M}{m}\dfrac{1}{2-B}$, as claimed. □

Therefore, if $M/m < 2-B$ as assumed in the statement, then the map $G(x)$ admits a unique fixed point $x^* \in [0,1]$, and hence the first-order condition (17) has a unique solution. Moreover, from (17) it immediately follows that $D_0^{\mathrm{W}} < \alpha/\underline{\theta}$, so that $\sqrt{2v} > \alpha/\underline{\theta}$ is a sufficient condition for $D_0^{\mathrm{W}} < \sqrt{2v}$.

We now turn to the cap D_1. The first-order condition with respect to D_1 is given by

$$\Xi(D_1) := (1-\gamma)\int_{\underline{\theta}}^{\theta_0(D_1)} (\alpha - \theta D_1)\,\mathrm{d}F(\theta) + \gamma \int_{\underline{\theta}}^{\overline{\theta}} (\alpha - \theta D_1)\,\mathrm{d}F_N(\theta) = 0, \tag{27}$$

which, rearranged, yields (18). The regulator's optimization problem is separable in D_0 and D_1 and

$$\Xi'(D_1) = -(1-\gamma)\left[\frac{\partial \theta_0}{\partial D_1}[\alpha - \theta_0(D_1)D_1]f(\theta_0) + F(\theta_0(D_1))\mathbb{E}_F[\theta|\theta \le \theta_0]\right] - \gamma \mathbb{E}_{F_N}[\theta] < 0,$$

because from the definition of $\theta_0(D_1)$ it follows that $\partial\theta_0/\partial D_1 > 0$ and

$$\alpha - \theta_0 D_1 = \left[\frac{F(\theta_0)}{f(\theta_0)} + \frac{\gamma}{(1-\gamma)f(\theta_0)}\right] D_1 > 0. \tag{28}$$

As W is globally concave in D_1, $D_1^{\mathrm{W}} < \widehat{D}$ if and only if $\Xi(\widehat{D}) < 0$, which is equivalent to

$$\mathbb{E}_{F_N}[\theta] > \frac{(\gamma + (1-\gamma)F(\theta_0))^2\theta_0}{\gamma(1-\gamma)F(\theta_0)} - \frac{(1-\gamma)F(\theta_0)\mathbb{E}_F[\theta|\theta \le \theta_0]}{\gamma}, \tag{29}$$

where θ_0 is the cutoff defined in (7). Therefore, assuming that v and $\mathbb{E}_{F_N}[\theta]$ are sufficiently large respectively guarantee that $D_0^{\mathrm{W}} < \sqrt{2v}$ and $D_1^{\mathrm{W}} < \widehat{D}$. Under these assumptions, social welfare is maximized at $(D_0^{\mathrm{W}}, D_1^{\mathrm{W}})$ if and only if $D_0^{\mathrm{W}} < D_1^{\mathrm{W}}$.

We next show that this inequality holds if and only if $\gamma < \gamma^*$, where $\gamma^* \in (0,1)$. Note that γ enters (17) only through its effect on $\theta_1(\cdot)$. By the Implicit Function Theorem and the second-order condition for the regulator's problem,

$$\mathrm{sgn}\left[\frac{\partial D_0^{\mathrm{W}}}{\partial \theta_1}\right] = \mathrm{sgn}\left[-(\alpha - \theta_1 D_0^{\mathrm{W}}) f(\theta_1)\right] < 0,$$

because $D_0^{\mathrm{W}} = \alpha / [\mathbb{E}_F\left[\theta | \theta \in [\theta_1(D_0^{\mathrm{W}}), \theta_2(D_0^{\mathrm{W}})]\right]] < \alpha/\theta_1(D_0^{\mathrm{W}})$. Since $\theta_1(\cdot)$ is decreasing in γ, we can conclude that D_0^{W} is increasing in γ. Since we have shown in the proof of Proposition 3 that $\widehat{D}$ is decreasing in γ, $\widehat{D} \to \alpha/\underline{\theta} > D_0^{\mathrm{W}}$ for $\gamma \to 0$ and $\widehat{D} \to 0 < D_0^{\mathrm{W}}$ for $\gamma \to 1$, we can conclude that there exists a threshold $\gamma^{**} \in (0,1)$ such that $D_0^{\mathrm{W}} < \widehat{D}$ if and only if $\gamma < \gamma^{**} \in (0,1)$.

Using the Implicit Function Theorem and the second-order condition for the regulator's problem, from (18) we have

$$\mathrm{sgn}\left[\frac{\partial D_1^{\mathrm{W}}}{\partial \gamma}\right] = \mathrm{sgn}\left[-\frac{\int_{\underline{\theta}}^{\theta_0(D_1^{\mathrm{W}})} (\alpha - \theta D_1^{\mathrm{W}})\,\mathrm{d}F(\theta)}{\gamma}\right] < 0,$$

because $\alpha - \theta_0 D_1^{\mathrm{W}} > 0$ from (28) and so $\alpha - \theta D_1^{\mathrm{W}} > 0$ for all $\theta \in [\underline{\theta}, \theta_0]$.

Suppose that (29) holds for all $\gamma < \gamma^{**} \in (0,1)$. Then, since $D_1^{\mathrm{W}} < \widehat{D}$ whenever $\widehat{D} > D_0^{\mathrm{W}}$, the above comparative statics results imply that there is a threshold $\gamma^* < \gamma^{**}$ such that, for all $\gamma < \gamma^*$, $\widehat{D} > D_1^{\mathrm{W}} > D_0^{\mathrm{W}}$, and so the considered solution maximizes welfare.

For $\gamma \geq \gamma^*$, the constraint $D_0 = D_1$ instead binds. Then, the welfare-optimal $D_0 = D_1 := D^{\mathrm{W}}$ is obtained solving

$$\max_{D^{\mathrm{W}}} \, (1-\gamma)\Bigg\{ v\mathbb{E}_F[\theta] + \int_{\underline{\theta}}^{\theta_2(D^{\mathrm{W}})} \left[\alpha D^{\mathrm{W}} - \frac{\theta}{2}(D^{\mathrm{W}})^2\right] \mathrm{d}F(\theta)$$

$$+ \int_{\theta_2(D^{\mathrm{W}})}^{\overline{\theta}} \left[\alpha D^{\mathrm{R}}(\theta) - \frac{\theta}{2}[D^{\mathrm{R}}(\theta)]^2\right] \mathrm{d}F(\theta)\Bigg\} + \gamma\left[\alpha D^{\mathrm{W}} - \mathbb{E}_{F_N}[\theta]\left(v - \frac{1}{2}(D^{\mathrm{W}})^2\right)\right].$$

The FOC of this problem is given by

$$(1-\gamma)\int_{\underline{\theta}}^{\theta_2(D^{\mathrm{W}})} (\alpha - \theta D^{\mathrm{W}})\,\mathrm{d}F(\theta) + \gamma \int_{\underline{\theta}}^{\overline{\theta}} (\alpha - \theta D^{\mathrm{W}})\,\mathrm{d}F_N(\theta) = 0,$$

and it can be written as in (19). As above, the definition of $\theta_2(\cdot)$ implies that the welfare objective is globally concave and that this equation admits a unique solution D^{W}, which is decreasing in γ. Since $D^{\mathrm{W}} < \alpha/\underline{\theta}$, $\sqrt{2v} > \alpha/\underline{\theta}$ again guarantees that $D^{\mathrm{W}} < \sqrt{2v}$. □

Proof of Lemma 1. Consumers' rents and incentive-compatibility constraints only depend on "net quality" levels $q(\theta) := v(\theta) - [D(\theta)]^2/2$. Fix any profile $\{q(\theta)\}_{\theta \in \Theta}$ that the platform may choose to offer. Suppose, by contradiction, that there is a type θ to whom the platform offers

$v(\theta) < \max_{\theta' \neq \theta} v(\theta')$. Holding fixed the net quality $q(\theta)$, the platform can extract a higher $D(\theta)$ (because we have assumed that the technological constraint does not bind) by raising $v(\theta)$ up to $\max_{\theta' \neq \theta} v(\theta')$. This is strictly profitable given that (1) it entails no additional quality-provision costs (as Ψ only depends on $\max_\theta v(\theta)$); (2) it increases revenues from data collection; and (3) it does not alter rents and incentive-compatibility constraints given that these only depend on net quality $q(\theta)$, which remains unaltered. This argument, which applies *vis-à-vis* sophisticated consumers, is unaffected by the presence of naive consumers that buy the cheapest option (provided $T \leq 0$) no matter v. Therefore, absent regulation the platform offers the same baseline service quality $\overline{v}$ to all consumers. □

Proof of Lemma 2. If the regulatory default option is offered at the full baseline service quality $\overline{v} := \max_\theta v(\theta)$, by the results of Lemma 1 the platform offers the same baseline service quality $\overline{v}^*$ to all consumers. Then, privacy-enhancing default regulation (i.e., $D_0 < \sqrt{2\overline{v}^*}$) introduces a binding constraint in the platform's screening problem *vis-à-vis* sophisticated consumers, implying that its profit is below the one it achieves under *laissez faire*, even if the platform can adjust its (uniform) baseline quality level. Offering instead the default option at a lower quality level, $v_0 < L(D_0)$ (where $v_0 < \overline{v}^*$ because $\overline{v}^* > L(D_0)$ for every privacy-enhancing default option), makes it unattractive to consumers, implying that the platform can still offer the *laissez faire* menu with $v = \overline{v}^*$ and so obtains a higher profit. □

Proof of Proposition 7. The arguments in Lemma 1 and 2 imply that the platform sets $v(\theta) := \overline{v}$ for all $\theta \in \Theta$, and w.l.o.g. for naive consumers as well, both under *laissez faire* and default regulation, provided that the regulator requires that D_0 shall be offered at the full service quality $\overline{v}$. Under *laissez faire*, the optimal quality level $\overline{v}^*$ is obtained solving

$$\max_v \, (1-\gamma)\Bigg[F(\theta_0)(\alpha\widehat{D} + \widehat{T}^*) + \int_{\theta_0}^{\theta_1} \left[\alpha D^*(\theta) + \left(\theta + \frac{F(\theta) - F(\theta_0)}{f(\theta)}\right)\left(v - \frac{1}{2}[D^*(\theta)]^2\right)\right] \mathrm{d}F(\theta)$$

$$\int_{\theta_1}^{\theta_2} \alpha\sqrt{2v}\,\mathrm{d}F(\theta) + \int_{\theta_2}^{\overline{\theta}} \left[\alpha D^*(\theta) + \left(\theta - \frac{1-F(\theta)}{f(\theta)}\right)\left(v - \frac{1}{2}[D^*(\theta)^2]\right)\right] \mathrm{d}F(\theta)\Bigg] + \gamma(\alpha\overline{D} + \widehat{T}^*) - \Psi(v),$$

where the data policies and cutoff types are those defined in Proposition (3), with $\widehat{D}$ denoting the bunching option for $\theta \leq \theta_0$, and $\widehat{T}^*$ the corresponding transfer. The marginal revenue of baseline service quality is thus given by

$$(1-\gamma)\left[\theta_1 F(\theta_0) + \int_{\theta_0}^{\theta_1}\left(\theta + \frac{F(\theta) - F(\theta_0)}{f(\theta)}\right)\mathrm{d}F(\theta) + \int_{\theta_1}^{\theta_2} \frac{2\alpha}{\sqrt{2v}}\,\mathrm{d}F(\theta) + \int_{\theta_2}^{\overline{\theta}}\left(\theta - \frac{1-F(\theta)}{f(\theta)}\right)\mathrm{d}F(\theta)\right] + \gamma\theta_1.$$

Under default regulation, the platform profit, gross of quality provision costs, is

$$(1-\gamma)\Bigg[F(\theta_0)(\alpha\widehat{D} + \widehat{T}^{\mathrm{R}}) + \int_{\theta_0}^{\theta_1(D_0)}\left[\alpha D^{\mathrm{R}}(\theta) + \left(\theta + \frac{F(\theta) - F(\theta_0)}{f(\theta)}\right)\left(v - \frac{1}{2}[D^{\mathrm{R}}(\theta)]^2\right) - U_0(\theta_1)\right]\mathrm{d}F(\theta)$$

$$+\int_{\theta_1(D_0)}^{\theta_2(D_0)} \alpha D_0 \,\mathrm{d}F(\theta)+\int_{\theta_2(D_0)}^{\overline{\theta}} \left[\alpha D^{\mathrm{R}}(\theta)+\left(\theta-\frac{1-F(\theta)}{f(\theta)}\right)\left(v-\frac{1}{2}[D^{\mathrm{R}}(\theta)^2]\right)-U_0(\theta_2)\right] \mathrm{d}F(\theta)\Bigg]+\gamma(\alpha\overline{D}+\widehat{T}^{\mathrm{R}}).$$

The data policies and cutoff types are those defined in Proposition (5), with $\widehat{D}$ denoting the bunching option for $\theta \leq \theta_0$, and $\widehat{T}^{\mathrm{R}}$ the corresponding transfer. These data policies and cutoffs do not depend on v. Therefore, the marginal revenue of baseline service quality equals

$$(1-\gamma)\left[\int_{\theta_0}^{\theta_1(D_0)}\left(\theta-\theta_1(D_0)+\frac{F(\theta)-F(\theta_0)}{f(\theta)}\right)\mathrm{d}F(\theta)+\int_{\theta_2(D_0)}^{\overline{\theta}}\left(\theta-\theta_2(D_0)-\frac{1-F(\theta)}{f(\theta)}\right)\mathrm{d}F(\theta)\right],$$

provided that consumer participation is ensured, i.e., $\overline{v}^{\mathrm{R}} \geq D_0^2/2$. Suppose that the regulator sets $D_0 = \sqrt{2\overline{v}^*}$. Comparing the expressions above immediately yields that the marginal benefit of baseline service quality at $\overline{v}^*$ absent regulation is higher than that of the regulated platform. This would imply that $\overline{v}^{\mathrm{R}} < \overline{v}^*$, but this choice would violate the participation constraint of consumers selecting the default option (as, by construction, $\overline{v}^{\mathrm{R}} - D_0^2/2 < \overline{v}^* - D_0^2/2 = 0$). Moreover, it can be checked that the derivative of the marginal benefit of quality for a regulated platform w.r.t. D_0 is identically zero, implying that the participation constraint would be violated for all $D_0 \leq \sqrt{2\overline{v}^*}$. As a result, whenever privacy-enhancing regulation is enacted, at the optimum the participation constraint of consumers in the pooling region must bind, i.e., $\overline{v}^{\mathrm{R}}(D_0) = D_0^2/2$. □

Proof of Proposition 8. Default regulation is inconsequential if $D_0 \geq \sqrt{2\overline{v}^*}$, as the default option is dominated by the free option $(\sqrt{2\overline{v}^*}, 0)$. The regulator thus solves

$$\max_{D_0 \leq \sqrt{2\overline{v}^*}} (1-\gamma)\int_{\underline{\theta}}^{\overline{\theta}}\left[\alpha D^{\mathrm{R}}(\theta)+\theta\left(\overline{v}^{\mathrm{R}}(D_0)-\frac{1}{2}[D^{\mathrm{R}}(\theta)]^2\right)\right]\mathrm{d}F(\theta)+\gamma\left[\alpha\overline{D}-\mathbb{E}_{F_N}[\theta]\left(\overline{v}^{\mathrm{R}}(D_0)-\frac{1}{2}\overline{D}^2\right)\right],$$

where $\overline{v}^{\mathrm{R}}(D_0) = D_0^2/2$. Taking the FOC and simplifying yields (20). Therefore, if binding default regulation is welfare-improving, the welfare-optimal default option solves this equation and is such that the SOC is satisfied, i.e., the LHS of (20) is decreasing at the considered stationary point. As $D = D_0^{\mathrm{W}}$ defined in (17) is such that the first term in (20) equals zero, it follows from the SOC that the welfare-optimal D_0 with endogenous quality is higher (resp. lower) than D_0^{W} if and only if the second term in (20) is positive (resp. negative) at $D = D_0^{\mathrm{W}}$, i.e., $\Psi'\left(\overline{v}^{\mathrm{R}}(D_0^{\mathrm{W}})\right) <$ (resp. ¿)$(1-\gamma)\mathbb{E}_F[\theta]+\gamma\mathbb{E}_{F_N}[\theta]$. Since the LHS of this inequality is the (increasing) marginal cost of baseline quality provision and the RHS the (constant) average consumers' valuation of this quality, this inequality is satisfied if and only if the platform would under-provide (resp. over-provide) quality from a social welfare standpoint when $D_0 = D_0^{\mathrm{W}}$. □

Proof of Proposition 9. We have argued in the text that, if the platform wants to extract $\overline{D}$ from naive consumers, it must offer $(\overline{D}, 0)$ as default option and $T^*(\theta) \geq 0$—hence, $D^*(\theta) \leq \sqrt{2v}$—to all sophisticated consumers. This constraint in the screening problem of the unregulated platform binds for $\theta \leq \theta_1$, implying that (by the concavity of the platform's objective in (21)) w.r.t. any

$D(\theta)$) all types $\theta \leq \theta_2$ are bunched at $(\sqrt{2v}, 0)$. The platform profit is then

$$\Pi_N(\gamma) := \gamma\alpha\overline{D} + (1-\gamma)\left[\int_{\underline{\theta}}^{\theta_2} \alpha\sqrt{2v}\,\mathrm{d}F(\theta) + \int_{\theta_2}^{\overline{\theta}} [\alpha D^*(\theta) + T^*(\theta)]\,\mathrm{d}F(\theta)\right],$$

where $(D^*(\theta), T^*(\theta))$ are given in (3) (because options with positive access prices are not affected by the presence of naive consumers).

Alternatively, the platform can make naive consumers aware of the privacy terms by offering non-default, negative-price options. Then, the optimal menu is that in Proposition 2, with the free option set as the default by the regulatory constraint. Because this menu features negative-price options under the considered parametric restriction $\underline{\theta} < \alpha/\sqrt{2v}$, naive consumers will become informed. The resulting platform profit, hereafter denoted by Π_S, is then independent of γ.

For $\gamma \geq \hat{\gamma}$, $\Pi_N(\gamma)$ equals the profit the platform makes absent informed consent regulation, because the platform offers the same menu in the two cases. This profit, $\Pi_N(\gamma|\gamma \geq \hat{\gamma})$, is larger than Π_S if a naive consumer is more profitable than a sophisticated consumer, even if the latter is offered the second-best menu in Proposition 2:

$$\overline{D} > \int_{\underline{\theta}}^{\overline{\theta}} \left[D^*(\theta) + \frac{T^*(\theta)}{\alpha}\right]\,\mathrm{d}F(\theta),$$

where $(D^*(\theta), T^*(\theta))$ are given in (3). This condition also implies that Π_N is increasing in γ. Moreover, $\Pi_N(0) < \Pi_S$ because the menu for sophisticated consumers in Π_N is inefficiently distorted away from the second-best menu in Proposition 2. Therefore, we can conclude that $\Pi_N(\gamma) < \Pi_S$ if and only if $\gamma < \tilde{\gamma}$, with $\tilde{\gamma} < \hat{\gamma}$. □

Proof of Proposition 10. If the regulator requires informed consent for consumers moving away from the default option and also mandates the platform to offer $(D_0, 0)$ as the default, with $D_0 \leq \sqrt{2v}$, the platform offers the menu in Proposition 5 with $\gamma = 0$. This is because the menu with $D_0 \leq \sqrt{2v}$ features $T^R(\theta) < 0$ for $\theta \leq \theta_1(D_0)$ (with $\theta_1(D_0) > \underline{\theta}$ implied by $\underline{\theta} < \alpha/\sqrt{2v}$), and so naive consumers will move away from the default and become sophisticated. Then, hard-cap regulation is never optimal and the welfare-optimal default option solves

$$\max_{D_0 \leq \sqrt{2v}} \int_{\underline{\theta}}^{\overline{\theta}} \left[\alpha D^{\mathrm{R}}(\theta) + \theta\left(v - \frac{1}{2}[D^{\mathrm{R}}(\theta)]^2\right)\right]\,\mathrm{d}F(\theta),$$

with $D^{\mathrm{R}}(\theta)$ given in (11) for $\gamma = 0$. Taking the FOC yields (17) for $\gamma = 0$. Under the uniform bounds assumption in Proposition 6 for $\gamma = 0$, we can conclude that the regulator optimally requires—on top of informed consent for consumers moving away from the default option—that the platform make the option $D_0 = \min\{\sqrt{2v}, D_0^{\mathrm{W}}|_{\gamma=0}\}$ available for free as the default. □

B Excerpts of selected EU laws

General Data Protection Regulation (GDPR)

Article 4

(1) "personal data" means any information relating to an identified or identifiable natural person ("data subject"); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

(2) "processing" means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction;

(3) "restriction of processing" means the marking of stored personal data with the aim of limiting their processing in the future;

(4) "profiling" means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements;

Article 6(1)

Processing shall be lawful only if and to the extent that at least one of the following applies:

(a) the data subject has given consent to the processing of his or her personal data for one or more specific purposes;

(b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;

(c) processing is necessary for compliance with a legal obligation to which the controller is subject;

(d) processing is necessary in order to protect the vital interests of the data subject or of another natural person;

(e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;

(f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.

Point (f) of the first subparagraph shall not apply to processing carried out by public authorities in the performance of their tasks.

Article 9

Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited.

Recital 32

Consent should be given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her, such as by a written statement, including by electronic means, or an oral statement. This could include ticking a box when visiting an internet website, choosing technical settings for information society services or another statement or conduct which clearly indicates in this context the data subject's acceptance of the proposed processing of his or her personal data. Silence, pre-ticked boxes or inactivity should not therefore constitute consent. Consent should cover all processing activities carried out for the same purpose or purposes. When the processing has multiple purposes, consent should be given for all of them. If the data subject's consent is to be given following a request by electronic means, the request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided.

The Digital Markets Act (DMA)

Recital 37

The less personalised alternative should not be different or of degraded quality compared to the service provided to the end users who provide consent, unless a degradation of quality is a direct consequence of the gatekeeper not being able to process such personal data or signing in end users to a service. Not giving consent should not be more difficult than giving consent. When the gatekeeper requests consent, it should proactively present a user-friendly solution to the end user to provide, modify or withdraw consent in an explicit, clear and straightforward manner. In particular, consent should be given by a clear affirmative action or statement establishing a freely given, specific, informed and unambiguous indication of agreement by the end user, as defined in Regulation (EU) 2016/679. At the time of giving consent, and only where applicable, the end user should be informed that not giving consent can lead to a less personalised offer, but that otherwise the core platform service will remain unchanged and that no functionalities will be suppressed. Exceptionally, if consent cannot be given directly to the gatekeeper's core platform service, end users should be able to give consent through each third-party service that makes use of that core platform service, to allow the gatekeeper to process personal data for the purposes of providing online advertising services.

Article 5

Obligations for gatekeepers

1. The gatekeeper shall comply with all obligations set out in this Article with respect to each of its core platform services listed in the designation decision pursuant to Article 3(9).

2. The gatekeeper shall not do any of the following:

   (a) process, for the purpose of providing online advertising services, personal data of end users using services of third parties that make use of core platform services of the gatekeeper;

   (b) combine personal data from the relevant core platform service with personal data from any further core platform services or from any other services provided by the gatekeeper or with personal data from third-party services;

   (c) cross-use personal data from the relevant core platform service in other services provided separately by the gatekeeper, including other core platform services, and vice versa; and

   (d) sign in end users to other services of the gatekeeper in order to combine personal data,

   unless the end user has been presented with the specific choice and has given consent within the meaning of Article 4, point (11), and Article 7 of Regulation (EU) 2016/679.

   Where the consent given for the purposes of the first subparagraph has been refused or withdrawn by the end user, the gatekeeper shall not repeat its request for consent for the same purpose more than once within a period of one year.

   This paragraph is without prejudice to the possibility for the gatekeeper to rely on Article 6(1), points (c), (d) and (e) of Regulation (EU) 2016/679, where applicable.

3. The gatekeeper shall not prevent business users from offering the same products or services to end users through third-party online intermediation services or through their own direct online sales channel at prices or conditions that are different from those offered through the online intermediation services of the gatekeeper.

4. The gatekeeper shall allow business users, free of charge, to communicate and promote offers, including under different conditions, to end users acquired via its core platform service or through other channels, and to conclude contracts with those end users, regardless of whether, for that purpose, they use the core platform services of the gatekeeper.

5. The gatekeeper shall allow end users to access and use, through its core platform services, content, subscriptions, features or other items, by using the software application of a business user, including where those end users acquired such items from the relevant business user without using the core platform services of the gatekeeper.

6. The gatekeeper shall not directly or indirectly prevent or restrict business users or end users from raising any issue of non-compliance with the relevant Union or national law by the gatekeeper with any relevant public authority, including national courts, related to any practice of the gatekeeper. This is without prejudice to the right of business users and gatekeepers to lay down in their agreements the terms of use of lawful complaints-handling mechanisms.

7. The gatekeeper shall not require end users to use, or business users to use, to offer, or to interoperate with, an identification service, a web browser engine or a payment service, or technical services that support the provision of payment services, such as payment systems for in-app purchases, of that gatekeeper in the context of services provided by the business users using that gatekeeper's core platform services.

8. The gatekeeper shall not require business users or end users to subscribe to, or register with, any further core platform services listed in the designation decision pursuant to Article 3(9) or which meet the thresholds in Article 3(2), point (b), as a condition for being able to use, access, sign up for or registering with any of that gatekeeper's core platform services listed pursuant to that Article.

9. The gatekeeper shall provide each advertiser to which it supplies online advertising services, or third parties authorised by advertisers, upon the advertiser's request, with information on a daily basis free of charge, concerning each advertisement placed by the advertiser, regarding:

   (a) the price and fees paid by that advertiser, including any deductions and surcharges, for each of the relevant online advertising services provided by the gatekeeper,

   (b) the remuneration received by the publisher, including any deductions and surcharges, subject to the publisher's consent; and

   (c) the metrics on which each of the prices, fees and remunerations are calculated.

   In the event that a publisher does not consent to the sharing of information regarding the remuneration received, as referred to in point (b) of the first subparagraph, the gatekeeper shall provide each advertiser free of charge with information concerning the daily average remuneration received by that publisher, including any deductions and surcharges, for the relevant advertisements.

10. The gatekeeper shall provide each publisher to which it supplies online advertising services, or third parties authorised by publishers, upon the publisher's request, with free of charge information on a daily basis, concerning each advertisement displayed on the publisher's inventory, regarding:

    (a) the remuneration received and the fees paid by that publisher, including any deductions and surcharges, for each of the relevant online advertising services provided by the gatekeeper;

(b) the price paid by the advertiser, including any deductions and surcharges, subject to the advertiser's consent; and

(c) the metrics on which each of the prices and remunerations are calculated.

In the event an advertiser does not consent to the sharing of information, the gatekeeper shall provide each publisher free of charge with information concerning the daily average price paid by that advertiser, including any deductions and surcharges, for the relevant advertisements.

11. The gatekeeper shall provide advertisers and publishers, as well as third parties authorised by advertisers and publishers, upon their request and free of charge, with access to the performance measuring tools of the gatekeeper and the data necessary for advertisers and publishers to carry out their own independent verification of the advertisements inventory, including aggregated and non-aggregated data. Such data shall be provided in a manner that enables advertisers and publishers to run their own verification and measurement tools to assess the performance of the core platform services provided for by the gatekeepers.

12. The gatekeeper shall provide end users and third parties authorised by an end user, at their request and free of charge, with effective portability of data provided by the end user or generated through the activity of the end user in the context of the use of the relevant core platform service, including by providing, free of charge, tools to facilitate the effective exercise of such data portability, and including by the provision of continuous and real-time access to such data.

13. The gatekeeper shall provide business users and third parties authorised by a business user, at their request, free of charge, with effective, high-quality, continuous and real-time access to, and use of, aggregated and non-aggregated data, including personal data, that is provided for or generated in the context of the use of the relevant core platform services or services provided together with, or in support of, the relevant core platform services by those business users and the end users engaging with the products or services provided by those business users. With regard to personal data, the gatekeeper shall provide for such access to, and use of, personal data only where the data are directly connected with the use effectuated by the end users in respect of the products or services offered by the relevant business user through the relevant core platform service, and when the end users opt in to such sharing by giving their consent.

14. The gatekeeper shall provide to any third-party undertaking providing online search engines, at its request, with access on fair, reasonable and non-discriminatory terms to ranking, query, click and view data in relation to free and paid search generated by end users on its online search engines. Any such query, click and view data that constitutes personal data shall be anonymised.

15. The gatekeeper shall apply fair, reasonable, and non-discriminatory general conditions of access for business users to its software application stores, online search engines and online social networking services listed in the designation decision pursuant to Article 3(9). For that purpose,

the gatekeeper shall publish general conditions of access, including an alternative dispute settlement mechanism. The Commission shall assess whether the published general conditions of access comply with this paragraph.

16. The gatekeeper shall not have general conditions for terminating the provision of a core platform service that are disproportionate. The gatekeeper shall ensure that the conditions of termination can be exercised without undue difficulty.

Article 6

Obligations for gatekeepers susceptible of being further specified under Article 8

1. The Gatekeeper shall comply with all obligations set out in this Article with respect to each of its core platform services listed in the designation decision pursuant to Article 3(9).

2. The gatekeeper shall not use, in competition with business users, any data that is not publicly available that is generated or provided by those business users in the context of their use of the relevant core platform services or of the services provided together with, or in support of, the relevant core platform services, including data generated or provided by the customers of those business users. For the purposes of the first subparagraph, the data that is not publicly available shall include any aggregated and non- aggregated data generated by business users that can be inferred from, or collected through, the commercial activities of business users or their customers, including click, search, view and voice data, on the relevant core platform services or on services provided together with, or in support of, the relevant core platform services of the gatekeeper.

3. The gatekeeper shall allow and technically enable end users to easily un-install any software applications on the operating system of the gatekeeper, without prejudice to the possibility for that gatekeeper to restrict such un-installation in relation to software applications that are essential for the functioning of the operating system or of the device and which cannot technically be offered on a standalone basis by third parties. The gatekeeper shall allow and technically enable end users to easily change default settings on the operating system, virtual assistant and web browser of the gatekeeper that direct or steer end users to products or services provided by the gatekeeper. That includes prompting end users, at the moment of the end users' first use of an online search engine, virtual assistant or web browser of the gatekeeper listed in the designation decision pursuant to Article 3(9), to choose, from a list of the main available service providers, the online search engine, virtual assistant or web browser to which the operating system of the gatekeeper directs or steers users by default, and the online search engine to which the virtual assistant and the web browser of the gatekeeper directs or steers users by default.

4. The gatekeeper shall allow and technically enable the installation and effective use of third-party software applications or software application stores using, or interoperating with, its operating

system and allow those software applications or software application stores to be accessed by means other than the relevant core platform services of that gatekeeper. The gatekeeper shall, where applicable, not prevent the downloaded third-party software applications or software application stores from prompting end users to decide whether they want to set that downloaded software application or software application store as their default. The gatekeeper shall technically enable end users who decide to set that downloaded software application or software application store as their default to carry out that change easily. The gatekeeper shall not be prevented from taking, to the extent that they are strictly necessary and proportionate, measures to ensure that third-party software applications or software application stores do not endanger the integrity of the hardware or operating system provided by the gatekeeper, provided that such measures are duly justified by the gatekeeper. Furthermore, the gatekeeper shall not be prevented from applying, to the extent that they are strictly necessary and proportionate, measures and settings other than default settings, enabling end users to effectively protect security in relation to third-party software applications or software application stores, provided that such measures and settings other than default settings are duly justified by the gatekeeper.

5. The gatekeeper shall not treat more favourably, in ranking and related indexing and crawling, services and products offered by the gatekeeper itself than similar services or products of a third party. The gatekeeper shall apply transparent, fair and non-discriminatory conditions to such ranking.

6. The gatekeeper shall not restrict technically or otherwise the ability of end users to switch between, and subscribe to, different software applications and services that are accessed using the core platform services of the gatekeeper, including as regards the choice of Internet access services for end users.

7. The gatekeeper shall allow providers of services and providers of hardware, free of charge, effective interoperability with, and access for the purposes of interoperability to, the same hardware and software features accessed or controlled via the operating system or virtual assistant listed in the designation decision pursuant to Article 3(9) as are available to services or hardware provided by the gatekeeper. Furthermore, the gatekeeper shall allow business users and alternative providers of services provided together with, or in support of, core platform services, free of charge, effective interoperability with, and access for the purposes of interoperability to, the same operating system, hardware or software features, regardless of whether those features are part of the operating system, as are available to, or used by, that gatekeeper when providing such services. The gatekeeper shall not be prevented from taking strictly necessary and proportionate measures to ensure that interoperability does not compromise the integrity of the operating system, virtual assistant, hardware or software features provided by the gatekeeper, provided that such measures are duly justified by the gatekeeper.

8. The gatekeeper shall provide advertisers and publishers, as well as third parties authorised by

advertisers and publishers, upon their request and free of charge, with access to the performance measuring tools of the gatekeeper and the data necessary for advertisers and publishers to carry out their own independent verification of the advertisements inventory, including aggregated and non-aggregated data. Such data shall be provided in a manner that enables advertisers and publishers to run their own verification and measurement tools to assess the performance of the core platform services provided for by the gatekeepers.

9. The gatekeeper shall provide end users and third parties authorised by an end user, at their request and free of charge, with effective portability of data provided by the end user or generated through the activity of the end user in the context of the use of the relevant core platform service, including by providing, free of charge, tools to facilitate the effective exercise of such data portability, and including by the provision of continuous and real-time access to such data.

10. The gatekeeper shall provide business users and third parties authorised by a business user, at their request, free of charge, with effective, high-quality, continuous and real-time access to, and use of, aggregated and non-aggregated data, including personal data, that is provided for or generated in the context of the use of the relevant core platform services or services provided together with, or in support of, the relevant core platform services by those business users and the end users engaging with the products or services provided by those business users. With regard to personal data, the gatekeeper shall provide for such access to, and use of, personal data only where the data are directly connected with the use effectuated by the end users in respect of the products or services offered by the relevant business user through the relevant core platform service, and when the end users opt in to such sharing by giving their consent.

11. The gatekeeper shall provide to any third-party undertaking providing online search engines, at its request, with access on fair, reasonable and non-discriminatory terms to ranking, query, click and view data in relation to free and paid search generated by end users on its online search engines. Any such query, click and view data that constitutes personal data shall be anonymised.

12. The gatekeeper shall apply fair, reasonable, and non-discriminatory general conditions of access for business users to its software application stores, online search engines and online social networking services listed in the designation decision pursuant to Article 3(9). For that purpose, the gatekeeper shall publish general conditions of access, including an alternative dispute settlement mechanism. The Commission shall assess whether the published general conditions of access comply with this paragraph.

13. The gatekeeper shall not have general conditions for terminating the provision of a core platform service that are disproportionate. The gatekeeper shall ensure that the conditions of termination can be exercised without undue difficulty.

The Digital Services Act (DSA)

Article 26

3. Providers of online platforms shall not present advertisements to recipients of the service based on profiling as defined in Article 4, point (4), of Regulation (EU) 2016/679 using special categories of personal data referred to in Article 9(1) of Regulation (EU) 2016/679.

References

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.

Acquisti, A., Brandimarte, L., and Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, 347(6221):509–514.

Acquisti, A., John, L. K., and Loewenstein, G. (2013). What is privacy worth? *The Journal of Legal Studies*, 42(2):249–274.

Acquisti, A., Taylor, C., and Wagman, L. (2016). The economics of privacy. *Journal of Economic Literature*, 54(2):442–492.

Anderson, S. P., Baik, A., and Larson, N. (2023). Price discrimination in the information age: Prices, poaching, and privacy with personalized targeted discounts. *The Review of Economic Studies*, 90(5):2085–2115.

Aridor, G., Che, Y.-K., and Salz, T. (2023). The effect of privacy regulation on the data industry: Empirical evidence from gdpr. *The RAND Journal of Economics*, 54(4):695–730.

Baron, D. P. and Myerson, R. B. (1982). Regulating a monopolist with unknown costs. *Econometrica*, 50(4):911–930.

Bergemann, D. and Bonatti, A. (2024). Data, competition, and digital platforms. *American Economic Review*, 114(8):2553–2595.

Bergemann, D., Bonatti, A., and Gan, T. (2022). The economics of social data. *The RAND Journal of Economics*, 53(2):263–296.

Besanko, D., Donnenfeld, S., and White, L. J. (1987). Monopoly and quality distortion: effects and remedies. *The Quarterly Journal of Economics*, 102(4):743–767.

Besanko, D., Donnenfeld, S., and White, L. J. (1988). The multiproduct firm, quality choice, and regulation. *The Journal of Industrial Economics*, 36(4):411–429.

Bisceglia, M. and Tirole, J. (2023). Fair gatekeeping in digital ecosystems. Working Paper 1452, Toulouse School of Economics.

Chen, J., Zhang, M., Guo, M., and Gao, Z. (2025). The role of risk attitudes in shaping digital privacy preferences: evidence from a large-scale survey. *Humanities and Social Sciences Communications*, 12(1):1–10.

Choi, J. P., Jeon, D.-S., and Kim, B.-C. (2019). Privacy and personal data collection with information externalities. *Journal of Public Economics*, 173:113–124.

Collis, A., Moehring, A., Sen, A., and Acquisti, A. (2021). Information frictions and heterogeneity in valuations of personal data. *Available at SSRN 3974826*.

Dworczak, P. and Muir, E. (2025). A mechanism-design approach to property rights. Available at SSRN 4637366.

Eilat, R., Eliaz, K., and Mu, X. (2021). Bayesian privacy. *Theoretical Economics*, 16(4):1557–1603.

Eliaz, K. and Spiegler, R. (2006). Contracting with diversely naive agents. *The Review of Economic Studies*, 73(3):689–714.

Eliaz, K. and Spiegler, R. (2008). Consumer optimism and price discrimination. *Theoretical Economics*, 3(4):459–497.

Farronato, C., Fradkin, A., and Lin, T. (2025). Designing consent: Choice architecture and consumer welfare in data sharing. Technical Report 34025, National Bureau of Economic Research.

Gabaix, X. and Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics*, 121(2):505–540.

Galperti, S., Liu, T., and Perego, J. (2024). Competitive markets for personal data. In *Proceedings of the ACM Conference on Economics and Computation (EC)*.

Gradwohl, R. (2018). Privacy in implementation. *Social Choice and Welfare*, 50(3):547–580.

Grande, D., Marti, X. L., Feuerstein-Simon, R., et al. (2020). Health policy and privacy challenges associated with digital technology. *JAMA Network Open*, 3(7):e208285.

Heidhues, P. and Kőszegi, B. (2017). Naivete-based discrimination. *The Quarterly Journal of Economics*, 132(2):1019–1054.

Ichihashi, S. (2020). Online privacy and information disclosure by consumers. *American Economic Review*, 110(2):569–595.

Johnen, J. (2019). Automatic-renewal contracts with heterogeneous consumer inertia. *Journal of Economics & Management Strategy*, 28(4):765–786.

Johnson, G. A. (2024). Economic research on privacy regulation: Lessons from the gdpr and beyond. In Goldfarb, A. and Tucker, C. E., editors, *The Economics of Privacy*. University of Chicago Press, Chicago, IL. Online edition, Chicago Scholarship Online, 22 May 2025, accessed 3 Oct. 2025.

Jullien, B. (2000). Participation constraints in adverse selection models. *Journal of Economic Theory*, 93(1):1–47.

Kirpalani, R. and Philippon, T. (2020). Data sharing and market power with two-sided platforms. Working Paper 28023, National Bureau of Economic Research.

Kőszegi, B. (2014). Behavioral contract theory. *Journal of Economic Literature*, 52(4):1075–1118.

Krähmer, D. and Strausz, R. (2023). Optimal nonlinear pricing with data-sensitive consumers. *American Economic Journal: Microeconomics*, 15(2):80–108.

Lewis, T. R. and Sappington, D. E. (1989). Countervailing incentives in agency problems. *Journal of Economic Theory*, 49(2):294–313.

Lin, T. (2022). Valuing intrinsic and instrumental preferences for privacy. *Marketing Science*, 41(4):663–681.

Loertscher, S. and Muir, E. (2025). Optimal hotelling auctions. Mimeo.

Maggi, G. and Rodríguez-Clare, A. (1995). On countervailing incentives. *Journal of Economic Theory*, 66(1):238–263.

Mertens, S., Herberz, M., Hahnel, U. J. J., and Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences of the United States of America*, 119(1):e2107346118.

Milgrom, P. and Shannon, C. (1994). Monotone comparative statics. *Econometrica: Journal of the Econometric Society*, pages 157–180.

Monti, G., Cremer, J., Fletcher, A., Heidhues, P., Jacobson, N., Kimmelman, G., and Schnitzer, M. (2025). Compliant use of personal data for advertising on social networks in europe. Policy Discussion Paper 11, Digital Regulation Project, Yale Tobin Center for Economic Policy.

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

Myerson, R. B. (1981). Optimal auction design. *Mathematics of operations research*, 6(1):58–73.

Pai, M. M. and Roth, A. (2013). Privacy and mechanism design. *ACM SIGecom Exchanges*, 12(1):8–29.

Sahni, N. S., Narayanan, S., and Kalyanam, K. (2019). An experimental investigation of the effects of retargeted advertising: The role of frequency and timing. *Journal of Marketing Research*, 56(3):401–418.

Seabrooke, L. and Wigan, D. (2014). Global wealth chains in the international political economy. *Review of International Political Economy*, 21(1):257–263.

Xiong, W., Xiong, Z., and Tian, T. (2022). Who to show the ad to? behavioral targeting in internet advertising. *Journal of Internet and Digital Economics*, 2(1):15–26.