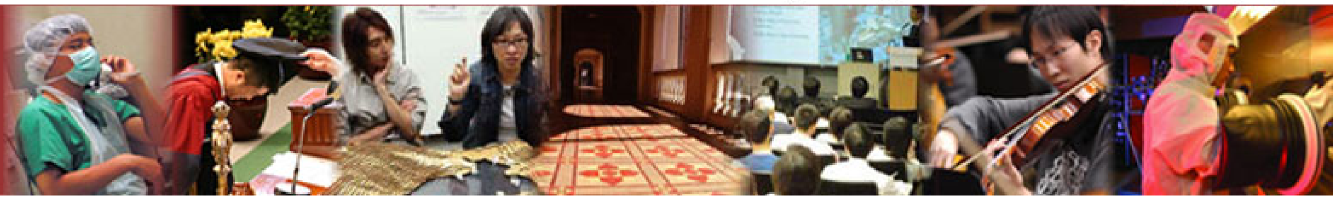
View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by HKU Scholars Hub

The HKU Scholars Hub The University of Hong Kong 香港大學學術庫



Title	Resolving the personalization-privacy dilemma: theory and implications of a privacy-preserving contract
Author(s)	Sin, RG; Jia, J
Citation	
Issued Date	2010
URL	http://hdl.handle.net/10722/161244
Rights	Creative Commons: Attribution 3.0 Hong Kong License

Resolving the Personalization-Privacy Dilemma: Theory and Implications of a Privacy-Preserving Contract

Raymond G. Sin and Jia Jia

School of Business & Management

Hong Kong University of Science & Technology

Clear Water Bay, Kowloon, Hong Kong

rsin@ust.hk

jiajia@ust.hk

Abstract

Personalization is an integral part of e-commerce strategy today. A unique feature of personalization is that it requires users to provide a certain amount of personal information to the service provider, thus giving rise to an interesting dilemma in that consumers cannot enjoy more personalized services without sacrificing more privacy. In this paper, we propose a mechanism that allows an online personalization vendor to provide proper incentives for consumers to share information, while protecting their privacy at the same time. The proposed solution not only enables consumers and the firm to engage in an otherwise unviable market, but it also allows the firm to implement an incentive-compatible menu that serves all consumers regardless of their privacy sensitivity. Further, we demonstrate that a minimum privacy-preservation policy is an effective device for protecting consumers' online privacy, and that it outperforms restricting vendors' ability in collecting customer information.

Our proposed mechanism is of theoretical and practical importance: By transforming the compensation schedule (privacy preservation) into a set-compliment device to the production variable, our approach offers an alternative to the reliance on external transfer, thus eradicating a major constraint confronted by traditional mechanism design. Practically, our research proposes a realistic, easily-implementable solution to the fervent calls for endowing consumers with greater control over their online privacy. Further, it offers important policy guidelines to the regulator on not only *what* devices can be applied in governing the information practice of online vendors, but also exactly *how* social-efficiency can be enhanced.

Keywords: personalization, privacy, mechanism design, welfare analysis
JEL Codes: L14, D82, L15, D61

1. Introduction

Personalization is an integral part of e-commerce strategy today. Web portals and online vendors offer a variety of personalization that ranges from browser helper objects (BHO) such as toolbars, to personalized recommendations, content deliveries, and search results (Hosanagar et al. 2008). Two prominent examples in web-based personalization are Amazon.com and Google. Amazon devotes over $200 million a year on investments in information technologies, of which a significant portion is dedicated for customer relationship management (CRM) and providing personalized offerings and recommendations through analyzing customer data. Google derives over 95% of their revenues from pure text-based advertisements, which are tailored to its users based on the specific queries that are submitted to Google's search engine. It also offers a variety of personalized services such as Google toolbar, Google Reader, and iGoogle. Recently, Google has expanded its venture to a new domain through the launch of Google Health, a web-based application that provides personalized health information management tool to its users.

Interestingly, even though the provision of personalization implies significant investments in technologies and infrastructures, most of these online services are provided free of charge to the consumers. The primary sources of revenue for vendors come typically not from charging for personalized services, but from the information that consumers divulge through the usage of these services. By assembling data from and forming detail profiles of their customers, firms can help third-party advertisers serve ads to the desired customer segments or use this information to launch their own targeted advertisements (e.g. Google, MSN, Yahoo!), share information with apps developers and claim part of their revenues (e.g. Apple and Facebook), harness the collective intelligence to offer products and services that better match customers' needs and induce repeat purchases (e.g. Amazon and Last.fm (Fleder and Hosanagar 2009)), or simply auction off the customer data (e.g. BlueKai (Steel 2009)).

A unique feature of personalization is that users can enjoy more personalization *only if* they are willing to disclose more information about themselves. For example, by entering a medical condition, a

Google Health user will be given reference links to learn about the symptoms, causes, and treatments about that particular condition. However, if the user would also like to know about potential interactions between the drugs she is taking and the allergies or conditions from which she is suffering, then additional information such as details of her medical records and prescriptions will also need to be provided. Therefore, personalization gives rise to an interesting dilemma in that consumers need to trade-off between enjoying personalized services and sacrificing privacy (Chellappa and Sin 2005; Awad and Krishnan 2006). Since the institutions gathering customer information also essentially "own" the information (Laudon 1996), consumers are often left with no control over how such information is used. The notion of potential abuse may trigger severe privacy concerns that deter consumers from adopting the services being offered (hence providing the corresponding information that is vital to online personalization vendors), or cause them to simply withdraw completely from participating in the market. Such dire consequences are exemplified in recent massive protests against Facebook's intrusive information practice, where some Facebook users boycotted certain services while others terminated their accounts. Therefore, privacy concerns may threaten not only the profitability of personalization providers, but also the very existence of the market itself. Our goal is to devise a mechanism that helps vendors to provide proper incentives for consumers to share information while protecting their privacy at the same time, and examine the welfare implications of regulating the vendor's information practice from a policy maker's perspective.

Our proposed mechanism is of theoretical and practical importance: Theoretically, we introduce information partition as a novel solution to the collapse in the number of available instruments that arises from the personalization-privacy tradeoff. By transforming the compensation schedule (privacy preservation) into a set-compliment device to the production variable, our design offers an alternative to the reliance on external instruments (such as monetary transfer), thus eradicating a major constraint confronted by traditional mechanism design. We demonstrate that not only does the optimal contract enable consumers and the firm to participate in an otherwise unviable market, but that by leaving consumers full control on part of the information acquired from them, the firm is also able to implement an incentive-compatible menu that serves the entire market.

Practically, our research proposes a realistic solution to the fervent calls for endowing consumers with greater control over their online privacy; interestingly, such voices are echoed not only among privacy advocates, but also the Internet giants including Google and Microsoft as well (Press 2008). The proposed contract can be implemented through new industry standards, such as the Platform for Privacy Preferences (P3P) protocol, that help consumers manage their privacy online by allowing them partial control on how much and in what way their information can be collected and used by online vendors. Further, our work is one of the first that responds to the Federal Trade Commission (FTC)'s initiatives in pursuing legislative options in protecting consumers' online privacy (FTC 2009; News 2010). Specifically, results from our welfare analysis offer important guidelines to the regulator on not only *what* devices can be applied in governing the information practice of online vendors, but more importantly, exactly *how* enhancements in social-efficiency can be achieved through a minimum privacy-preservation policy. Our research has far-reaching implications, especially given the astonishing growth in the delivery of personalized digital contents for mobile devices that marks the inception of ubiquitous personalization in the digital era.

This paper proceeds as follows: Section 2 reviews relevant literature on personalization and mechanism design, which is followed by the development of the basic model in section 3. In section 4, we examine vendor's optimal strategy under asymmetric information, and discuss the welfare implications of alternative regulatory interventions. Section 5 concludes the paper with discussions on theoretical and managerial implications of our work.

2. Literature Review

The first formal study on consumers' tradeoffs between personalization and privacy in the online context was conducted by Chellappa and Sin (2005). In their study, the authors empirically examine the respective roles of consumers' valuation of personalized services and the privacy costs associated with sharing personal information in their likelihood of adopting online personalization. Their research establishes that values for personalization and concerns for privacy are two independent factors, yet they jointly determine a consumer's decision to use personalized services. Further, individuals may experience different levels of privacy concerns even in disclosing the same amount of information. Such systematic differences in the preference for privacy, which is referred to as "privacy sensitivity" and is typically unobservable to a third-party (Chellappa and Shivendu 2010), may be attributed to individuals defining information spaces of their social lives differently, being exposed to different situational cues and primed differently with the consequences, or placing different probabilities or values on a given outcome (Hann et al. 2007).

Mechanism design has been recommended as a useful framework in studying the market for personal information, where firms can device a mechanism that allows consumers to "purchase" an ideal level of privacy (Rust et al. 2002; Murthi and Sarkar 2003). Ever since the seminal works by Mussa and Rosen (1978) and Maskin and Riley (1984) on monopolistic nonlinear pricing, mechanism design has been applied to examine a broad variety of issues, such as governmental regulation of the private monopolist (Baron and Myerson 1982), procurement (Laffont and Tirole 1993), and duopolistic competition in horizontally and vertically differentiated markets (Rochet and Stole 2002). In addition to the rich literature in economics on both theoretical development and applications of this framework, there has recently been a few applications in Information Systems on topics ranging from versioning and pricing of information goods (Sundararajan 2004; Huang and Sundararajan forthcoming 2011), to optimal sampling strategy and digital rights management (Sundararajan 2004). In particular, Chellappa and Shivendu (2010) are among the first to embrace this approach in studying the market for personalization, where the amount and quality of personalization that can be offered to a customer is proportional to the amount of information that the firm has about her (Volokh 2000; Adomavicius and Tuzhilin 2002). The authors propose couponing as a compensation instrument in designing a segmenting mechanism to address consumers' intrinsic privacy concerns associated with sharing personal information. Despite the novelty of their modeling approach, the applicability of such side payments or monetary subsidies in the market for personalization is limited in practice. The closest real-life examples are Amazon's "A9 Instant Reward" and Microsoft's "Bing cashback" programs [1]. Unfortunately, both programs suffered similar fates in being discontinued within two years of their introduction (Turner and Wolfson 2010); most online vendors nowadays still rely primarily on offering personalized services as incentives for consumers to share information.

Although privacy is a complicated concept and has been termed many different definitions in the literatures of sociology, law, and political science (Solove 2008), in this research we shall focus on the informational aspect of privacy (Stone et al. 1983; Stone and Stone 1990), which is of utmost concern to online customers and is most relevant in the context of personalization (Culnan and Armstrong 1999). Specifically, we consider the loss of control over the outflow and subsequent usage of one's own information as the source of privacy costs that arise from sharing personal information. We propose "privacy preservation", which involves granting consumers complete control over part of the information collected from them by the online vendor, as a solution to the personalization-privacy dilemma. The idea that such preservation can moderate consumers' privacy concern is supported by extant literature. In particular, Hann et al. (2007) find that proper information handling and access procedures can mitigate privacy concerns and result in an increase in willingness to provide personal information. Tsai et al. (forthcoming 2011) suggest that individuals are willing to pay a premium for privacy once privacy information is made more prominent and intuitive. Interestingly, even for mechanisms that are non-verifiable, such as privacy

---

[1] A9 Instant Reward was more popularly known as the "$\pi/2$ discount", which offered monetary rewards to users of their browser-embedded toolbar that delivers personalized search results. This program was launched in 2004 and discontinued in 2006. Bing cashback was launched in 2008 and discontinued in 2010.

statements or seals, their presence still mitigates privacy concerns by helping consumers assess of the risks of disclosing personal information to websites (Milne and Culnan 2004; Hui et al. 2007). These findings provide strong evidence that by delivering a sense of control over their own information, firms can manage consumers' privacy concern and strategically leverage privacy protection for competitive advantages.

3. Model

We consider a market where consumers are heterogeneous in their privacy sensitivity, denoted by $\theta \in [\underline{\theta}, \overline{\theta}]$, such that consumers with higher values of θ are more privacy sensitive. We assume θ to be private information to the consumers, and is distributed with density function $f(\theta)$ and cumulative density $F(\theta)$; where $f(\theta)$ is continuously differentiable, everywhere positive, and log-concave on its support $[\underline{\theta}, \overline{\theta}]$.[2]

A principal — an online portal or vendor — offers free personalization services that are considered valuable by the consumers. Consumers provide personal information that is required to enjoy the corresponding personalized services. Once this information (denoted by I) is acquired and processed, a spectrum of personalized offerings is being generated in return. The value (also referred to as convenience hereafter) that a consumer derives from consuming these services (denoted by $S(I)$) is assumed to be positively correlated with the amount of information provided as well as the extent to which the vendor is

[2] If $f(\theta)$ is continuously differentiable and log-concave, then it is single-peaked (uni-modal), and the reciprocal of its hazard rate, i.e., $\frac{F(\theta)}{f(\theta)}$, is non-decreasing, satisfying the monotone hazard rate property. Most common distributions satisfy these standard assumptions (Bagnoli and Bergstrom 2005). Further, with log-concavity of $f(\theta)$, $F(\theta)$ is also log-concave on $[\underline{\theta}, \overline{\theta}]$ (a standard proof is provided in the online appendix).

able to use such information in generating personalization. For tractability, we use a linear function to capture this intrinsic relationship between convenience and information acquisition:

$$S(I) = aI \tag{3.1}$$

where the parameter a captures the vendor's technological efficiency in generating convenience for the consumers from their personal information. This relationship is assumed to be deterministic, and is common knowledge to all parties in the market. Similar treatments can be found in Chellappa and Shivendu (2010).

Consumers face the personalization-privacy dilemma documented in extant literature: one cannot enjoy more convenience (personalized services) without sacrificing more privacy (personal information). Hence, a consumer's utility increases with amount of personalization consumed and decreases with the amount of personal information shared. The following function captures this tradeoff:

$$U(I,\theta) = S(I) - \theta I \tag{3.2}$$

Note that in the absence of a contract that protects consumers' privacy, consumers have no control over the information that has been transferred to the vendor. Hence the disutility that arises from a consumer's privacy concern is assumed to be proportional to the full set of information provided to the vendor, scaled by her privacy sensitivity θ.

Given $S = aI$, equation (3.2) can be rewritten as a function of I; i.e. $U(I,\theta) = (a-\theta)I$. Hence a consumer's decision on whether to subscribe to the proposed personalization services (or to "participate" in the market) depends on the relative magnitude of the efficiency coefficient a against her type coefficient θ. We label those consumers with $\theta > a$ "privacy seekers", as the costs of their privacy concerns outweigh the benefits derived from personalization so that they abstain from participating in the market; and those with $\theta \leq a$ "convenience seekers", as these consumers value convenience so much so that they are willing disclose as much information as possible in exchange for personalization. It can be observed that when the market comprises of only privacy seekers (i.e. $\underline{\theta} > a$), no consumer is willing to participate given the negative utility associated with using any level of personalization. In other words, the personalization market is unviable. Our goal is to provide a solution under such circumstances by introducing the notion of privacy preservation.

Privacy preservation (η) is defined as the portion of I (the set of information acquired by the vendor) that is used purely for the purpose of generating personalization. It is ex-ante contractible (e.g. through p3p or conventional privacy policies) and ex-post verifiable and enforceable (e.g. through auditing and government sanctions). With the credible commitment that η will not be subject to secondary use, consumers' privacy concerns arise only from the disclosure of its complement (i). Formally,

$$i = I - \eta \tag{3.3}$$

Hence under the preservation scheme, the utility that a consumer of type θ derives upon adopting personalization is given by:

$$U(I, \eta, \theta) = S(I) - \theta i \tag{3.4}$$

3.1 Vendor's objective

The online vendor generates revenue from the set of information over which it can explore commercial possibilities, and incurs a constant marginal cost[3] in converting the acquired information into personalized services to its consumers. Without loss of generality, the marginal cost is normalized to 1. This normalization suggests that the magnitudes of other coefficients in our model are to be interpreted in a relative sense, rather than an absolute one. The vendor's profit from serving each consumer is defined as:

$$\pi(I, \eta) = \varphi(i) - I \tag{3.5}$$

[3] This can be interpreted as a "resource cost" – the cost associated with the necessary computing resources in providing personalized content to a request (Liu et al. 2010).

where $\varphi(i)$ is the revenue-generating function of the vendor, and is assumed to be of the following form:

$$\varphi(i) = bi - \frac{1}{2}i^2 \tag{3.6}$$

$b \in \mathbb{R}^+$ in equation (3.6) denotes the commercial efficiency of the vendor's personalization strategy; i.e. a larger b implies more effective use of consumers' information in generating revenues. The quadratic functional form is chosen for analytical convenience and is common in mechanism design literature (Rochet and Stole 2002; Yang and Ye 2008).

Due to information asymmetry of consumers' privacy sensitivity (θ), the vendor's problem is to design a direct mechanism of $\{I(\theta), \eta(\theta)\}_{\theta \in [\underline{\theta}, \bar{\theta}]}$ so that consumers of different types self-select into the desirable pair, allowing the vendor to maximize profit. Formally, the vendor's programme is:

$$\begin{aligned}
&\int_{\underline{\theta}}^{\bar{\theta}} \left\{ b\big(I(\theta) - \eta(\theta)\big) - \frac{1}{2}\big(I(\theta) - \eta(\theta)\big)^2 - I(\theta) \right\} f(\theta) d\theta \\
&s.t. \quad \text{for each } \theta, \qquad \theta = \arg\max_{\tilde{\theta}} aI(\tilde{\theta}) - \theta(I(\tilde{\theta}) - \eta(\tilde{\theta})) \qquad \text{(I.C.)} \\
&\qquad\qquad\qquad\qquad aI(\theta) - \theta\big(I(\theta) - \eta(\theta)\big) \geq 0 \qquad \text{(I.R.)}
\end{aligned} \tag{3.7}$$

Our model formulation is unique in that: 1) the two contracting variables in our model are additive separable, since our instruments are set-compliment (i.e. η is a subset of I)[4]; 2) the compensatory transfer is perceived differently by the principal and the agents; i.e. the principal's cost of delivering services is not identical to the benefits received by the agent. Despite these departures from the conventional mechanism design problem, our model is analogous to the traditional setup in two fundamental ways.

[4] Therefore, besides the participation and incentive compatibility constraints, an additional constraint $I(\theta) \geq \eta(\theta)$ needs to be accounted for when evaluating the implementability of the contract. Technical details are available from Lemma A1 in the Online Appendix.

First, the contracting pair $\{I(\theta),\eta(\theta)\}$ satisfies the single-crossing property (i.e. $(\partial/\partial\theta)(\partial_{\eta}U/\partial_{I}U)>0 \quad \forall\theta$ in consumer's utility $aI(\theta)-\theta(I(\theta)-\eta(\theta))$). Hence local analysis is sufficient to ensure global incentive compatibility. Second, both the production and compensation functions are separable and well-defined. To see this, we can rewrite the vendor's marginal gain from serving customers of type θ as $bi(\theta)-\frac{1}{2}i(\theta)^{2}-(i(\theta)+\eta(\theta))$ (from equations (3.3) and (3.6)) and consumer's utility as $a\eta(\theta)-(\theta-a)i(\theta)$ (from Equations (3.1) and (3.4)). $i(\theta)$ can be interpreted as the production variable for both the principal and the agent, while $\eta(\theta)$ assumes the role of the compensation variable. Under complete information, maximizing profit necessarily implies that the vendor keeps $\eta(\theta)$ at the lowest possible level, which is determined by the participation constraint $a\eta(\theta)-(\theta-a)i(\theta)\geq 0$. For a given production level $i(\theta)$, the vendor acquires and processes extra information, defined by $\eta(\theta)=\frac{(\theta-a)}{a}i(\theta)$, in order to offset the production cost $(\theta-a)i(\theta)$ borne by the consumer. Hence the tradeoff faced by the vendor reduces to $bi(\theta)-\frac{1}{2}i(\theta)^{2}-\left(i(\theta)+\frac{(\theta-a)}{a}i(\theta)\right)$, and that the compensation affects vendor's decision on the optimal production level in a similar fashion as in traditional models.

4. Analysis

4.1 Optimality of the privacy-preserving contract

In this section, we characterize the optimal contract that is composed of information acquisition and privacy preservation. We show that the proposed information partition is not only incentive-compatible, but it also allows the vendor to fully separate the market. We shall first present some preliminary results based on complete information before proceeding to the asymmetric information case.

Lemma 1. Under complete information, the optimal allocation $\{I^*(\theta),\eta^*(\theta)\}$ fulfills $I^*(\theta)=\frac{\theta}{a}\left(b-\frac{\theta}{a}\right)$ and $\eta^*(\theta)=\frac{(\theta-a)}{a}\left(b-\frac{\theta}{a}\right)$.

All proofs are relegated to the Appendix. We can observe from Lemma 1 that the production variable is given by $i^*(\theta)=I^*(\theta)-\eta^*(\theta)=b-\frac{\theta}{a}$, which is decreasing in θ. This is consistent with the intuition that more efficient types (less privacy-sensitive consumers in our case) produce more (allow more information to be used for commercial purposes). On the other hand, $I^{*}{}'(\theta)>0$ implies that more information is being extracted from more privacy-sensitive consumers. While this result appears counter-intuitive at the first glance, it should be noted that the preservation level also increases in θ, and that it increases at a faster rate compared to information acquisition (i.e. $\eta^{*}{}'(\theta)>I^{*}{}'(\theta)>0$). Together they suggest that more values are indeed being delivered to the highly privacy-sensitive consumers to ensure their participation; because an increasing portion of the additional information being gathered is used to generate personalized services for these consumers rather than revenues for the firm.

The following lemma shows that under asymmetric information, the infinite incentive constraints reduce to a simple differential equation with a monotonicity condition. This lemma helps us narrow down the search for an optimal profile among the implementable ones.

Lemma 2: For $\underline{\theta}>a$, a piecewise C^1 incentive-compatible allocation $\{I(\theta),\eta(\theta)\}$ satisfies $d\eta/d\theta\le 0$ and $(a-\theta)\frac{dI}{d\theta}+\theta\frac{d\eta}{d\theta}=0$ a.e. on $[\underline{\theta},\overline{\theta}]$.

Lemma 2 implies that production level decreases with consumers' privacy sensitivity $\left(\frac{di}{d\theta}=\frac{a}{\theta-a}\frac{d\eta}{d\theta}\le 0\right)$, which is analogous to the monotonicity condition in standard mechanism design problems. Since reporting a higher θ than their actual type allows consumers to enjoy a saving in production cost, incentive compatibility of the objective menu requires that the preservation level be lower for more privacy-sensitive consumers: the vendor associates a $\frac{\theta - a}{a}$ decrease in preservation level with a marginal decrease in production level; together they result in a reduction in the level of information acquisition by $\frac{\theta}{a}$. These dynamics are being captured by the differential equation. Hence, contrary to the complete information case presented in Lemma 1, both information acquisition and privacy preservation are non-increasing in consumer's privacy sensitivity.

Having established Lemma 2, we can now formally characterize the optimal implementable menu in the following proposition.

Proposition 1. The optimal allocation $\{\tilde{I}(\theta), \tilde{\eta}(\theta)\}$ is fully-separating, and is characterized by the following system of equations:

$$\begin{cases} \tilde{\eta}(\theta) = \frac{1}{a}\left\{(\theta - a)\left(b - \frac{1}{a}\left(\theta + \frac{F(\theta)}{f(\theta)}\right)\right) + \int_{\theta}^{\bar{\theta}} \left(b - \frac{1}{a}\left(v + \frac{F(v)}{f(v)}\right)\right) dv\right\} \\ \tilde{I}(\theta) = \frac{1}{a}\left\{\theta\left(b - \frac{1}{a}\left(\theta + \frac{F(\theta)}{f(\theta)}\right)\right) + \int_{\theta}^{\bar{\theta}} \left(b - \frac{1}{a}\left(v + \frac{F(v)}{f(v)}\right)\right) dv\right\} \end{cases} \text{on } [\underline{\theta}, \bar{\theta}].$$

We can observe from Proposition 1 that a unique contract-pair is being offered to consumers of each type (full separation). Further, the production level delegated to a consumer of type θ is $\tilde{i}(\theta) = b - \frac{1}{a}\left(\theta + \frac{F(\theta)}{f(\theta)}\right)$. Compared with that under complete information, the vendor faces a downward adjustment in the production level by $\frac{1}{a}\frac{F(\theta)}{f(\theta)}$. This reduction follows from incentive compatibility of the contract, which dictates that the vendor leaves some surplus $\left(-U_{\theta}\left(I(\theta), \eta(\theta), \theta\right) = i(\theta)\right)$ – also known as the information rent – to a consumer of marginally more efficient type to prevent her from deviating. Since the vendor needs to process $\frac{i(\theta)}{a}$ amount of additional information to generate this surplus, the total cost associated with delivering the rent to those with lower privacy sensitivity than type θ consumers is $\frac{1}{a}\int_{\underline{\theta}}^{\theta} i(\theta) f(v)dv = \frac{i(\theta)F(\theta)}{a}$. The underproduction is a consequence of accounting for this additional cost as the vendor determines $\tilde{i}(\theta)$.

The integral component in the contract, $\int_{\theta}^{\bar{\theta}} \tilde{i}(v)dv = \int_{\theta}^{\bar{\theta}} \left(b - \frac{1}{a}\left(v + \frac{F(v)}{f(v)}\right)\right)dv$, is the amount of information rent required to prevent a consumer of type θ from pretending to be more privacy sensitive than she is. It is derived from aggregating incremental rents $\left(\tilde{i}(\theta)\right)$ from the most privacy-sensitive consumers, and is delivered by the vendor through additional preservations.

Figure 1 presents a graphical illustration of the optimal contract under asymmetric information, and demonstrates how the slopes of the two contracting variables $\left\{\tilde{I}(\theta), \tilde{\eta}(\theta)\right\}$ change with respect to the technological efficiency coefficient (a). It can be observed that both information acquisition and privacy preservation approach zero when technological efficiency is sufficiently low. This implies that shutting down a portion of the market (intentionally excluding consumers of certain types) may indeed be more profitable for the vendor under certain circumstances. The following proposition characterizes the optimal market coverage from the vendor's perspective.

Proposition 2. The optimal market coverage of $\left\{\tilde{I}(\theta), \tilde{\eta}(\theta)\right\}$ is defined by $[\underline{\theta}, \theta^*]$, where $\theta^* = \min\{\bar{\theta}, \theta^s\}$ and θ^s is the solution of $ab = \left(\theta + \frac{F(\theta)}{f(\theta)}\right)$.

Proposition 2 shows that market coverage is increasing in both technological and commercial efficiencies of the vendor $\left(\frac{\partial \theta^s}{\partial a} > 0 \text{ and } \frac{\partial \theta^s}{\partial b} > 0\right)$, and that the whole market is served as long as $ab \geq \left(\bar{\theta} + \frac{F(\bar{\theta})}{f(\bar{\theta})}\right)$. Intuitively, the more effectively the vendor generates revenues from commercially using consumer information (i.e., the larger b is), the more eager he is to acquire such information and expand the market. The positive association between personalization efficiency (a) and market coverage, however, is less apparent. From the privacy-preservation level $\tilde{\eta}(\theta)=\frac{1}{a}\left\{(\theta-a)\tilde{i}(\theta)+\int_{\theta}^{\bar{\theta}}\tilde{i}(v)dv\right\}$, we can observe that an improvement in technological efficiency not only reduces consumers' marginal production cost $(\theta-a)$ – which by itself diminishes the shutdown portion as consumers find it less costly to participate – but also magnifies the compensatory role of $\tilde{\eta}(\theta)$; i.e. for the same level of information to be used for commercial purposes $(i(\theta))$, less privacy-preservation is required. Hence the vendor is motivated to serve a larger market while using a larger portion of the acquired information for generating revenues. For the remaining analyses, we shall assume b to be sufficiently large to induce full market coverage.

We have now characterized the optimal solution, with the premise that the vendor has complete freedom in specifying both the levels of information acquisition and privacy preservation in the contract. In the following subsections, we relax this assumption and consider the scenario where a regulator takes a more proactive role in protecting consumers' privacy online. Specifically, we investigate the welfare implications of two legislative options, namely the minimum privacy-preservation policy and the maximum information-acquisition policy. Our goal is to inform policy makers on the choice of an optimal regulatory device and the corresponding magnitude of control in inducing efficient market outcomes.

4.2 Minimum privacy-preservation policy

Under a minimum privacy-preservation policy, the vendor is mandated to reserve a predetermined proportion of information (also referred to as "preservation ratio") from any secondary use for all consumers that he serves. Our solution approach involves a two-stage process: In the first stage, the regulator sets a preserving ratio; then conditional on this additional constraint, the vendor designs the optimal contract in the second stage. We derive the optimal solution from the regulator's perspective using backward induction by first considering the effects of a given preservation ratio on the vendor's objective. Formally, we introduce the following inequality into the vendor's programme[5]:

$$\frac{\eta(\theta)}{I(\theta)} \geq \alpha \tag{4.1}$$

where $1 > \alpha \geq \underline{\alpha}$. $\underline{\alpha} = \frac{\tilde{\eta}(\underline{\theta})}{\tilde{I}(\underline{\theta})}$ is lowest attainable preservation in a contract based solely on the vendor's self-interest $\left(\frac{d}{d\theta}\left(\frac{\tilde{\eta}(\theta)}{\tilde{I}(\theta)}\right) \geq 0\right)$; hence it defines the lower bound of an effective regulatory intervention.

The following lemma establishes that given this additional preservation requirement, the vendor finds that a partially-bunching solution dominates their fully-separating counterparts. Further, it indicates that solutions with more than one bunching interval or with bunching intervals at other locations cannot be optimal for the vendor.

Lemma 3. The optimal menu $\left\{I^P(\theta), \eta^P(\theta)\right\}$ bunches the low-end of the type space $[\underline{\theta}, \hat{\theta}]$.

This lemma leads to the following proposition.

Proposition 3. The optimal contract under a given preservation ratio is characterized by the followings:

1. There exists a lower interval $[\underline{\theta}, \hat{\theta}]$ where bunching takes place. $\hat{\theta}$ is jointly determined by the following two equations:

$$\int_{\underline{\theta}}^{\hat{\theta}} \left\{ \left(\left(\theta + \frac{F(\theta)}{f(\theta)} \right) - \left(\hat{\theta} + \frac{F(\hat{\theta}) - (1-\alpha)\lambda}{f(\hat{\theta})} \right) \right) f(\theta) \right\} d\theta + \left(a - \hat{\theta}(1-\alpha) \right) \lambda = 0 \tag{4.2}$$

[5] The reason behind expressing the minimum preservation as a ratio to the total amount of information acquired by the vendor, as opposed to an absolute value, is that otherwise the vendor can respond by arbitrarily increasing information acquisition to counteract the effects of such a policy.

and $(1-\alpha)\iota(\hat{\theta},\lambda,\alpha)=b-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\theta')}\right)$ (4.3)

where $\iota(\theta,\lambda,\alpha)=\frac{1}{a}\left\{\theta\left(b-\frac{1}{a}\left(\theta+\frac{F(\theta)-(1-\alpha)\lambda}{f(\theta)}\right)\right)+\int_{\theta}^{\bar{\theta}}\left(b-\frac{1}{a}\left(v+\frac{F(v)-(1-\alpha)\lambda}{f(v)}\right)\right)dv\right\}$.

2. In the upper interval of the type space $[\hat{\theta},\bar{\theta}]$, the optimal solution $\{I^P(\theta,\lambda,\alpha),\eta^P(\theta,\lambda,\alpha)\}$ satisfies

$$\begin{cases} I^P(\theta,\lambda,\alpha)=\frac{1}{a}\left\{\theta\left(b-\frac{1}{a}\left(\theta+\frac{F(\theta)-(1-\alpha)\lambda}{f(\theta)}\right)\right)+\int_{\theta}^{\bar{\theta}}\left(b-\frac{1}{a}\left(v+\frac{F(v)-(1-\alpha)\lambda}{f(v)}\right)\right)dv\right\} \\ \eta^P(\theta,\lambda,\alpha)=\frac{1}{a}\left\{(\theta-a)\left(b-\frac{1}{a}\left(\theta+\frac{F(\theta)-(1-\alpha)\lambda}{f(\theta)}\right)\right)+\int_{\theta}^{\bar{\theta}}\left(b-\frac{1}{a}\left(v+\frac{F(v)-(1-\alpha)\lambda}{f(v)}\right)\right)dv\right\} \end{cases},$$

whereas in the lower interval of the type space $[\underline{\theta},\hat{\theta}]$, it satisfies $\begin{cases} I^P(\theta,\lambda,\alpha)=I^P(\hat{\theta},\lambda,\alpha) \\ \eta^P(\theta,\lambda,\alpha)=\eta^P(\hat{\theta},\lambda,\alpha) \end{cases}$ and

$\eta^P(\hat{\theta},\lambda,\alpha)=\alpha I^P(\hat{\theta},\lambda,\alpha)$.

3. $F(\theta)-(1-\alpha)\lambda \ge F(\hat{\theta})-(1-\alpha)\lambda \ge 0$ for $\theta\in[\hat{\theta},\bar{\theta}]$.

In all of the above equations, λ is the shadow price associated with the inequality (4.1). It ensures

$\frac{\partial}{\partial\theta}\left(\theta+\frac{F(\theta)-(1-\alpha)\lambda}{f(\theta)}\right)\ge 0$ on $[\hat{\theta},\bar{\theta}]$.

The first two parts of the proposition describe how the vendor can best respond to a mandated level of minimum preservation ratio in the contract. It indicates that simply bunching the segment with ratios smaller than α cannot be an optimal solution. Instead, a superior strategy for the vendor is to also revise production levels by increasing the production level by $\frac{(1-\alpha)\lambda}{f(\theta)}$ for all consumers falling under the separating region. The last part of this proposition indicates that the underproduction problem for the less efficient consumers ($\theta\in[\hat{\theta},\bar{\theta}]$) are moderated compared with that in the benchmark model, yet not significant enough to induce overproduction. As a result, both consumer surplus and vendor's profit increase on that segment.

For consumers falling under the bunching region ($\theta \in [\underline{\theta}, \hat{\theta}]$), productions are uniformly set to the level of $b - \frac{1}{a}\left(\hat{\theta} + \frac{F(\hat{\theta}) - (1-\alpha)\lambda}{f(\hat{\theta})}\right)$. Compared with those in the benchmark case $\left(b - \frac{1}{a}\left(\theta + \frac{F(\theta)}{f(\theta)}\right)\right)$, the underproduction problem is more significant for consumers with sufficiently low privacy sensitivities while less severe for those who are in the neighborhood of $\hat{\theta}$. The overall effects of bunching the less privacy-sensitive consumers $\left(\theta \in [\underline{\theta}, \hat{\theta}]\right)$ is a suppression of the aggregate production $\left(\frac{1}{a}\int_{\underline{\theta}}^{\hat{\theta}}\left(\left(\theta + \frac{F(\theta)}{f(\theta)}\right) - \left(\hat{\theta} + \frac{F(\hat{\theta}) - (1-\alpha)\lambda}{f(\hat{\theta})}\right)\right) f(\theta) d\theta > 0\right)$. On the other hand, this suppression also leads to a relaxation of the minimum-preservation constraint for the separating region where underproduction is alleviated. The associated gain is captured by $\frac{(1-\alpha)}{a}\left(\frac{a}{(1-\alpha)} - \hat{\theta}\right)\lambda$, which should be equal to the loss in production efficiency on the bunching region on optimality. This tradeoff is captured by equation (4.2).

The shape of the optimal menu under uniform distribution and a given regulation level is illustrated in Figure 2. We can observe that though a minimum privacy-preservation policy results in the vendor suppressing the production of the lower-end segments of the market; it also enables the vendor to increase production in the remaining segments. Figure 3 contrasts the preservation ratio under the regulatory solution with that in the benchmark case, and demonstrates the ratio to be increasing in α in the separating region, while being uniformly larger than that of the baseline menu except at $\overline{\theta}$.

Based on equations (4.2) and (4.3) , the following corollary establishes how the magnitude of the ratio affects the shape of the optimal contract.

Corollary 1. $\hat{\theta}'(\alpha) > 0$: the vendor bunches a larger portion of the market when faced with a higher preservation ratio. There exists a threshold $(\bar{\alpha})$ beyond which the vendor serves all consumers with an identical contract $\left\{ I^P(\alpha) = \frac{1}{(1-\alpha)}\left(b - \frac{1}{(1-\alpha)}\right), \eta^P(\alpha) = \frac{\alpha}{(1-\alpha)}\left(b - \frac{1}{(1-\alpha)}\right) \right\}$.

Having established corollary 1, we can now formulate the social planner's objective as a piecewise function of α on $[0,1)$:

$$SW(\alpha) = \begin{cases} \int_{\underline{\theta}}^{\bar{\theta}} \left\{ b\left(\tilde{I}(\theta) - \tilde{\eta}(\theta)\right) - \frac{1}{2}\left(\tilde{I}(\theta) - \tilde{\eta}(\theta)\right)^2 - \frac{\theta}{a}\left(\tilde{I}(\theta) - \tilde{\eta}(\theta)\right) \right\} f(\theta) d\theta & \forall \alpha \in [0, \underline{\alpha}) \\ \int_{\underline{\theta}}^{\hat{\theta}(\alpha)} \left\{ b\left(i^P\left(\hat{\theta}(\alpha), \lambda(\alpha), \alpha\right)\right) - \frac{1}{2} i^P\left(\hat{\theta}(\alpha), \lambda(\alpha), \alpha\right)^2 \right. & \\ \quad \left. - \frac{\theta}{a}\left(i^P\left(\hat{\theta}(\alpha), \lambda(\alpha), \alpha\right)\right) \right\} f(\theta) d\theta + \int_{\hat{\theta}(\alpha)}^{\bar{\theta}} \left\{ b\left(i^P\left(\theta, \lambda(\alpha), \alpha\right)\right) \right. & \forall \alpha \in [\underline{\alpha}, \bar{\alpha}) \\ \quad \left. - \frac{1}{2}\left(i^P\left(\theta, \lambda(\alpha), \alpha\right)\right)^2 - \frac{\theta}{a}\left(i^P\left(\theta, \lambda(\alpha), \alpha\right)\right) \right\} f(\theta) d\theta & \\ \int_{\underline{\theta}}^{\bar{\theta}} \left\{ b\left((1-\alpha) I^P(\alpha)\right) - \frac{1}{2}\left((1-\alpha) I^P(\alpha)\right)^2 \right. & \\ \quad \left. - \frac{\theta}{a}(1-\alpha) I^P(\alpha) \right\} f(\theta) d\theta & \forall \alpha \in [\underline{\alpha}, \bar{\alpha}) \end{cases} \tag{4.4}$$

The following proposition establishes that regulatory intervention by imposing a minimum preservation ratio improves social welfare. It also characterizes the optimal degree of intervention from the social-planner's perspective.

Proposition 4. A minimum privacy-preservation policy is welfare-enhancing; in particular, the optimal regulation level lies within $\alpha \in (\underline{\alpha}, \bar{\alpha})$, inducing a partially-separating market outcome.

Proposition 4 offers important insights for a policy maker pursuing legislative options in governing the information practice of online personalization vendors. Our findings suggest that although the vendor has incentives to introduce privacy-preservation in the contract design, the level of this self-serving preservation is not sufficient from the society's perspective. We demonstrate that social efficiency can be enhanced by simply introducing a minimum preservation policy by which the vendor is required to preserve a portion of the acquired information from any secondary uses. The fact that the optimal minimum preservation level lies strictly above the attainable preservation in the absence of intervention demonstrates the policy to be an invaluable device for the regulator.

Further, we find that the social planner prefers to induce a partially-separating market outcome. This is a counter-intuitive finding, as one may expect that a fully-separating menu that serves each consumer with a different degree of information acquisition and privacy preservation to be of best interest to the society (an equivalence of a benevolent provider of a public goods charging more from the rich and less from the poor). In fact, this is precisely the case under the complete information, when the vendor can extract full surplus from consumers having exactly the same objective as that of the social planner. Under asymmetric information, however, the lack of information renders full extraction no longer possible. The distorted incentive leads the vendor to adopt a more aggressive differentiation strategy than is required to achieve social-efficiency, resulting in severe underproduction from the societal perspective. By imposing a minimum privacy-preservation policy, the regulator constrains the vendor's ability to exercise excessive discrimination yet not to the extent that completely deprives him of the ability to differentiate (i.e. when $\alpha = \overline{\alpha}$).

Figure 4 presents a graphical illustration of social welfare under different preservation levels with uniform distribution. We can observe that the social welfare peaks at a point strictly smaller than $\overline{\alpha}$. The kink that we observe at $\overline{\alpha}$ reflects the fact that participation constraint no longer binds for $\alpha \in (\overline{\alpha}, 1)$.

4.3 Maximum information-acquisition policy

We now explore an alternative regulation whereby the policy maker restricts the vendor's ability in gathering customer information by setting an upper bound on the amount of information to be acquired. We refer to this upper bound as an "information boundary". Similar to the analysis on the minimum privacy-preservation policy, our solution approach here also involves a two-stage process with backward induction. We shall first characterize the vendor's optimal contract in the existence of an information boundary. Formally, we introduce the following inequality into the vendor's programme:

$$I(\theta) \le K < \tilde{I}(\underline{\theta}) \tag{4.5}$$

where $K < \tilde{I}(\underline{\theta})$ is a positive finite number, and ensures that the information boundary constraint is relevant for consumers with certain degrees of privacy sensitivity. The following lemma shows that full separation is no longer optimal.

Lemma 4. When $K < \tilde{I}(\underline{\theta})$, the optimal menu $\left\{I^B(\theta), \eta^B(\theta)\right\}$ bunches the low-end of the type space $[\underline{\theta}, \hat{\theta}]$.

This lemma serves the same role as Lemma 3 does for the previous subsection, and leads to the following proposition.

Proposition 5. The optimal contract under an information boundary is characterized by the following conditions:

1. There exists a lower interval $[\underline{\theta}, \hat{\theta}]$ that corresponds to the mass of consumers who are being served an identical contract, while the remaining market is served with a separating menu. $\hat{\theta}$ is jointly determined by the following two equations:

$$\int_{\underline{\theta}}^{\hat{\theta}} \left\{ \left(\theta + \frac{F(\theta)}{f(\theta)} \right) - \left(\hat{\theta} + \frac{F(\hat{\theta}) + \lambda}{f(\hat{\theta})} \right) \right\} f(\theta) d\theta + \hat{\theta}\lambda = 0 \tag{4.6}$$

$$\text{and } \kappa(\hat{\theta}, \lambda) = K \tag{4.7}$$

where $\kappa(\theta, \lambda) = \frac{\theta}{a}\left(b - \frac{1}{a}\left(\theta + \frac{F(\theta) + \lambda}{f(\theta)} \right) \right) + \frac{1}{a}\int_{\theta}^{\bar{\theta}} \left(b - \frac{1}{a}\left(v + \frac{F(v) + \lambda}{f(v)} \right) \right) dv$.

2. In the upper interval of the type space $[\hat{\theta},\bar{\theta}]$, the optimal solution $\left\{I^B(\theta,\lambda),\eta^B(\theta,\lambda)\right\}$ satisfies

$$\begin{cases} I^B(\theta,\lambda)=\dfrac{1}{a}\left\{\theta\left(b-\dfrac{1}{a}\left(\theta+\dfrac{F(\theta)+\lambda}{f(\theta)}\right)\right)+\int_{\theta}^{\bar{\theta}}\left(b-\dfrac{1}{a}\left(v+\dfrac{F(v)+\lambda}{f(v)}\right)\right)dv\right\} \\ \eta^B(\theta,\lambda)=\dfrac{1}{a}\left\{(\theta-a)\left(b-\dfrac{1}{a}\left(\theta+\dfrac{F(\theta)+\lambda}{f(\theta)}\right)\right)+\int_{\theta}^{\bar{\theta}}\left(b-\dfrac{1}{a}\left(v+\dfrac{F(v)+\lambda}{f(v)}\right)\right)dv\right\} \end{cases},$$

whereas in the lower interval of the type space $[\underline{\theta},\hat{\theta}]$, it satisfies $\begin{cases} I^B(\theta,\lambda)=I^B(\hat{\theta},\lambda) \\ \eta^B(\theta,\lambda)=\eta^B(\hat{\theta},\lambda) \end{cases}$.

In the above equations, λ is the shadow value associated with the information boundary constraint (equation (4.5)); it ensures that $\dfrac{\partial}{\partial\theta}\left(\theta+\dfrac{F(\theta)+\lambda}{f(\theta)}\right)\geq 0$ on $[\hat{\theta},\bar{\theta}]$.

The shadow value measures the marginal increase in revenue with a uniform relaxation of the information boundary on the bunching area $\left([\underline{\theta},\hat{\theta}]\right)$. We can observe from Proposition 5 that λ is also effective in determining the contracts for those whose information boundary is not relevant. The underlying reason is that a marginal increase in information acquisition for an unconstrained type of consumers (who lie in the separating interval) creates an externality for all their constrained peers (who fall on the bunching interval) through a chain reaction dictated by the monotonicity condition in Lemma 2. As a result, the tightening of information boundary for constrained consumers leads to an increase in the shadow price that the vendor needs to account for when designing contracts for consumers falling within the unconstrained interval.

We can also observe that the existence of an information boundary lowers both information acquisition and privacy preservation for all consumers compared with the second-best solution presented in Proposition 1. Further, this reduction is more severe for consumers with lower privacy sensitivity $\left(\left(\tilde{I}(\theta)-I^B(\theta,\lambda)\right)_\theta\leq 0,\left(\tilde{\eta}(\theta)-\eta^B(\theta,\lambda)\right)_\theta\leq 0\right)$; the reason is that the marginal payoffs from utilizing their information are relatively low from the vendor's perspective when they already produce at high levels.

The fact that we observe a bunching solution for the most efficient segment of the market appears to contradict the conventional understanding in mechanism design, which suggests that the vendor always prefers to differentiate the most efficient portion of the market (e.g., Mussa and Rosen, 1978); i.e. the vendor would bunch (and require lower production from) consumers with high or intermediate privacy-sensitivities in exchange for the ability to differentiate consumers with low privacy-sensitivity. However, Proposition 5 indicates that these alternatives are not feasible profit-maximizing strategies for two reasons.

First, the monotonicity condition that is automatically fulfilled for the most efficient consumers in the baseline model[6] may be violated in the presence of an information boundary. The introduction of an information boundary reduces the slackness of the monotonicity condition for the whole market, causing information acquisition for the less privacy-sensitive consumers to diminish at a faster rate $\left(\left(\tilde{I}(\theta)-I^{B}(\theta)\right)_{\lambda\theta}<0\right)$; hence the monotonicity condition starts to bind from the lower end of the market $(\underline{\theta})$. Bunching those consumers is the only remedy for retaining the monotonicity condition and thus the incentive-compatibility of the contractible menu.

Second, bunching the most efficient types allows the vendor to increase the production efficiency of the less efficient consumers. To illustrate the underlying dynamics, we construct a fully-separating menu (thereby suffice to ensure the fulfillment of the monotonicity condition) that serves the high-types, i.e. $[\hat{\theta},\overline{\theta}]$ ($\hat{\theta}$ is the point at which the menu reaches the information cap (K)), while being incentive compatibility on the target segment. Equation (4.7) describes the method of constructing such a menu.

[6] This can be generalized to any distribution with bounded support. Particularly, Barlow et al. (1963) show that for a distribution with its support bounded from above, the inverse hazard rate $\left(\frac{1-F(\theta)}{f(\theta)}\right)$ is non-increasing in $(\overline{\theta}-\varepsilon,\overline{\theta}]$. Analogously, we can prove the dual case; that is, for a distribution with its support bounded from below, $\frac{F(\theta)}{f(\theta)}$ is non-decreasing in $[\underline{\theta},\underline{\theta}+\varepsilon)$.

$\frac{\partial \lambda}{\partial \hat{\theta}} < 0$ derived from equation (4.7) implies that with a fixed information boundary, serving a larger market with a fully-separating menu induces more severe distortions in production ($i(\theta,\lambda) = b - \frac{1}{a}\left(\theta + \frac{F(\theta)+\lambda}{f(\theta)}\right)$) on $[\hat{\theta},\overline{\theta}]$, an effect that is economically equivalent to lowering the information boundary. Narrowing the market coverage of this separating menu (shifting $\hat{\theta}$ to the right) thus moderates the underproduction problem, but at the cost of suppressing productions from the remaining segment at type $\hat{\theta}$'s level. Proposition 5 establishes that the gain from the former $\left(\hat{\theta}\lambda\right)$ should balances the loss from the latter $\left(\int_{\underline{\theta}}^{\hat{\theta}}\left\{\left(\hat{\theta}+\frac{F(\hat{\theta})+\lambda}{f(\hat{\theta})}\right)-\left(\theta+\frac{F(\theta)}{f(\theta)}\right)\right\} f(\theta)d\theta\right)$ on optimality, which is captured by equation (4.6).

To illustrate the general results derived in Proposition 5, we apply a uniform distribution to consumer types, and graphically represent the optimal menus under different levels of information boundary in Figure 5. It can be observed that the bunching regions of the optimal menu expand with lower information boundaries. Such an expansion continues until the boundary reaches a certain threshold, beyond which a purely-bunching solution consistently outperforms any menu of other structures (note that this result does not hold for more general distributions; see corollary 2 for details). In addition, the lowering of information boundary results in more severe underproduction problem.

Given that $\hat{\theta}$ and λ in equations (4.6) and (4.7) are both endogenous in K, we use $\hat{\theta}(K)$ and $\lambda(K)$ to denote these two functions. The following corollary summarizes how the vendor's optimal bunching region changes in response to different levels of information boundary.

Corollary 2.

1. $\hat{\theta}'(K) < 0$: lowering information boundary leads to an expansion of the bunching interval.

2. $\hat{\theta}(K) \in [\underline{\theta}, \min\{\tilde{\theta}, \overline{\theta}\})$ where $\tilde{\theta} \equiv \inf\left\{\theta : \theta - \frac{F(\theta)}{f(\theta)} = 0\right\}$: the optimal menu does not necessarily degenerate to a purely-bunching one even under a sufficiently low information boundary.

The expansion of the bunching interval is straight forward. The second result, however, is somewhat counter-intuitive. Specifically, it suggests that if there exists some θ such that $\theta - \frac{F(\theta)}{f(\theta)} \leq 0$, then differentiating the least efficient segment of the market is always a profit-maximization strategy. Further, only in a market with consumer types being evenly distributed $\left(\theta > \frac{F(\theta)}{f(\theta)}\right)$ everywhere on its support can a purely-bunching menu emerge.

Given the monotonicity of $\hat{\theta}(K)$ established in corollary 2, we can now formulate the social planner's objective as a function of K :

For the set of partially-separating menus,

$$SW(K) = \int_{\underline{\theta}}^{\hat{\theta}(K)} \left(b\left(I^B(\hat{\theta},\lambda) - \eta^B(\hat{\theta},\lambda)\right) - \frac{1}{2}\left(I^B(\hat{\theta},\lambda) - \eta^B(\hat{\theta},\lambda)\right)^2 - \frac{\theta}{a}\left(I^B(\hat{\theta},\lambda) - \eta^B(\hat{\theta},\lambda)\right)\right) f(\theta)d\theta$$
$$+\int_{\hat{\theta}(K)}^{\overline{\theta}} \left(b\left(I^B(\theta,\lambda) - \eta^B(\theta,\lambda)\right) - \frac{1}{2}\left(I^B(\theta,\lambda) - \eta^B(\theta,\lambda)\right)^2 - \frac{\theta}{a}\left(I^B(\theta,\lambda) - \eta^B(\theta,\lambda)\right)\right) f(\theta)d\theta$$

while for the set of purely-bunching menus,

$$SW(K) = \int_{\underline{\theta}}^{\overline{\theta}} \left(b\left(I^B(K) - \eta^B(K)\right) - \frac{1}{2}\left(I^B(K) - \eta^B(K)\right)^2 - \frac{\theta}{a}\left(I^B(K) - \eta^B(K)\right)\right) f(\theta)d\theta$$

where $I^B(K) = K$ and $\eta^B(K) = \frac{\overline{\theta} - a}{\overline{\theta}} K$ from the fact that the participation constraint binds at $\overline{\theta}$. The result is summarized in the following proposition.

Proposition 6. A maximum information-acquisition policy results in welfare reduction.

The result presented in proposition 6 stands in stark contrast with that under a minimum-privacy-preservation policy, which *improves* social welfare. The intuition behind this result is that an information boundary first takes effects on the least privacy-sensitive consumers, who might demand more personalization than is available under this policy, while without offering any effective assistance to the more privacy-sensitive consumers. Hence an information boundary leads to an unambiguous reduction in both consumer and vendor surpluses.

5. Conclusion

5.1 Implications of our work

Information technology plays a critical role in both the provision and usage of personalization by online vendors and consumers. On one hand, more sophisticated technologies allow firms to deliver more and better personalized services to consumers while seamlessly sharing more information with their associates, thus allowing for higher profit potentials; on the other hand, the very same technologies that enable firms to share larger amounts of information more quickly with more parties also exacerbate consumers' concerns for privacy. Recent controversies surrounding Facebook's "Instant Personalization" program is a vivid example. Under Instant Personalization, Facebook shares users' information with their associates, allowing them to "instantly" personalize the web experience for those users even if it were their first time interacting with the websites. Facebook's information practice has not only prompted public uproars and protests from various consumer groups, but also resulted in some Facebook users boycotting certain services while others terminating their accounts altogether (Anonymous 2010). Therefore, the tradeoffs between personalization and privacy give rise to an interesting predicament: without the ability to use customer information to generate revenues, firms would not have incentives to offer free personalization; without a mechanism in place to govern the usage and protection of customer data, consumers would not share their personal information. Our research investigates this seemingly evident yet subtle issue from the perspectives of both consumers and online personalization provider.

Our answer to the personalization-privacy dilemma is to transform the original compensation problem into an information partition one. We show that the vendor can design a contractible menu using privacy-preservation as an instrument, which allows consumers with varying privacy sensitivity to self-select into adopting a desirable duo of information acquisition and protection. As a result, full participation can be induced even under the most restrictive market conditions, and a second-best outcome is achieved when the exact levels of privacy sensitivity of individual consumers are unknown to the vendor.

The contributions of our work are two-fold: Theoretically, our technique for deriving an optimal implementable contract in the absence of any external instruments is ground-breaking in mechanism design. To our best knowledge, we are the first to propose a contract-theoretic framework that eradicates the reliance on external transfer common in traditional designs. While readers may notice certain similarities between our approach and the revenue/cost sharing rules established in extant literature, two attributes of our model define its uniqueness: first, existing models rely primarily on monetary instruments, while ours is applicable also to non-price contexts; second, revenues sacrificed by the principal contribute to a *reduction in cost*, as opposed to an increment in transfer or benefits, to the agent. Our methodology thus offers a systematic approach in solving a new class of adverse selection problems in which production is intrinsically correlated with compensation; for example, it can be applied to investigate the development of an open source project that depends solely on voluntary contribution, while the management board's objective is to maximize the overall efforts from all participants. Though contribution is (heterogeneously) costly to developers, higher effort levels are intrinsically correlated with larger personal payoffs; e.g. being more fun, gaining deeper knowledge, and performance improvements in their paid work (e.g. Lerner and Tirole 2005). Practically, not only does our proposed mechanism provide a solution to an otherwise unviable market, but that the contract is also incentive-compatible and can easily be implemented in reality. Our research offers important guidelines to online personalization vendors and policy makers in addressing the growing concerns over online privacy.

First, our analyses on preservation ratio and information boundary offer valuable insights and concrete recommendations to online personalization vendors regarding how an optimal contract can be revised when confronted with different limitations. Specifically, we show that the vendor can benefit from increasing the bunching interval, making an identical offer to more consumers when the preservation ratio is sufficiently high, while increasing the production level by the less efficient segment. On the other hand, the optimal strategy for a vendor confronting a tight information boundary is to bunch the efficient segment while reducing the production level by the most privacy-sensitive consumers at the same time.

Second, while online personalization vendors may have economic incentives to provide consumers with preservation through some form of privacy-management settings, the level of protection delivered by these options may not be sufficient from the regulator's perspective (Giles 2010). Our welfare analysis suggests that a minimum privacy-preservation policy can not only serve as a simple yet powerful device, but also a superior legislative option compared with limiting vendors' ability in acquiring customer information in inducing a more socially-efficient outcome.

5.2 Future research directions

In this paper, we seek to provide a solution to the case where consumers' privacy concerns are sufficiently high to threaten the de-facto existence of the personalization market. Our proposed mechanism can be readily extended to examine a more general scenario with the coexistence of both privacy-seekers, who refrain from participating completely, and convenience-seeking consumers who prefer as much personalization as possible. Further, a sensitivity analysis on the distribution of consumer types under the generalized market conditions can reveal the effects of changes in consumers' privacy concerns on the optimal contract and equilibrium market outcome. Finally, we have assumed in our model that technology efficiency is exogenously determined; in reality, online portals often need to decide on the investment level in personalization technologies prior to engaging in an information-sharing contract with the consumers. A model that endogenizes such an investment decision and hence the corresponding efficiency by which the vendor is able to generate values to their consumers would be an interesting direction to pursue.

Reference

Adomavicius, G. and A. Tuzhilin. 2002. An Architecture of e-Butler- A Consumer Centric Online Personalization Mechanism. *International Journal of Computational Intelligence and Applications* **3**(2)313-327.

Anonymous. 2010. *Privacy 2.0*. *The Economist*. **394** 18.

Awad, N. F. and M. S. Krishnan. 2006. The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to be Profiled Online for Personalization. *MIS Quarterly* **30**(1)13-28.

Bagnoli, M. and T. Bergstrom. 2005. Log-concave probability and its applications. *Economic Theory* **26**(2)445-469.

Barlow, R. E., A. W. Marshall and F. Proschan. 1963. Properties of Probability Distributions with Monotone Hazard Rate. *The Annals of Mathematical Statistics* **34**(2)375-389.

Baron, D. P. and R. B. Myerson. 1982. Regulating a Monopolist with Unknown Costs. *Econometrica* **50**(4)911-930.

Chellappa, R. K. and S. Shivendu. 2010. Mechanism Design for "Free" but "No Free Disposal" Services: The Economics of Personalization Under Privacy Concerns. *Management Science* **56**(10)1766-1780.

Chellappa, R. K. and R. Sin. 2005. Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma. *Information Technology and Management* **6**(2-3)181-202.

Culnan, M. J. and P. K. Armstrong. 1999. Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science* **10**(1)104-115.

FTC. 2009. *Beyond Voice: Mapping the Mobile Marketplace* Federal Trade Commission.

Giles, M. 2010. Privacy 2.0 - Give a little, take a little. *Economist* **394**(8667)18-19.

Hann, I.-H., K.-L. Hui, S.-Y. Lee and I. Png. 2007. Overcoming Online Information Privacy Concerns: An Information-Processing Theory Approach. *Journal of Management Information Systems* **24**(2)13-42.

Huang, K.-W. and A. Sundararajan. forthcoming 2011. Pricing Digital Goods: Discontinuous Costs and Shared Infrastructure. *Information Systems Research*.

Hui, K.-L., H. H. Teo and S.-Y. T. Lee. 2007. The Value of Privacy Assurance: An Exploratory Field Experiment. *MIS Quarterly* **31**(1)19-33.

Laffont, J.-J. and J. Tirole. 1993. *A Theory of Incentives in Procurement and Regulation*. The MIT Press, Cambridge.

Laudon, K. C. 1996. Markets and privacy. *Communications of The ACM* **39**(9)92-104.

Lerner, J. and J. Tirole. 2005. The Economics of Technology Sharing: Open Source and Beyond. *Journal of Economic Perspectives* **19**(2)99-120.

Liu, D., S. Sarkar and C. Sriskandarajah. 2010. Resource Allocation Policies for Personalization in Content Delivery Sites. *Information Systems Research* **21**(2)227-248.

Maskin, E. and J. T. Riley. 1984. Monopoly with Incomplete Information. *The RAND Journal of Economics* **15**(2)171-196.

Milne, G. R. and M. J. Culnan. 2004. Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices. *Journal of Interactive Marketing* **18**(3)15-29.

Murthi, B. P. S. and S. Sarkar. 2003. The Role of the Management Sciences in Research on Personalization. *Management Science* **49**(10)1344-1362.

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

News, F. 2010. *FTC Testifies on Protecting Teen Privacy*.

Press, A. 2008. *Microsoft, Google Demand Privacy Legislation*.

Rochet, J. and L. A. T. Stole. 2002. Nonlinear Pricing with Random Participation. *The Review of Economic Studies* **69**(1)277-311.

Rust, R. T., P. K. Kannan and N. Peng. 2002. The Customer Economics of Internet Privacy. *Journal of the Academy of Marketing Science* **30**(4)455-464.

Seierstad, A. and K. Sydsæter. 1987. *Optimal control theory with economic applications*. North-Holland.

Solove, D. J. 2008. *Understanding Privacy*. Harvard University Press.

Steel, E. 2009. *Target-Marketing Becomes More Communal*. *The Wall Street Journal* B10.

Stone, E. F., D. G. Gardner, H. G. Gueutal and s. McCLure. 1983. A Field Experiment Comparing Information-Privacy Values, Beliefs, and Attitudes Across Several Types of Organizations. *Journal of Applied Psychology* **68**(3)459-468.

Stone, E. F. and D. L. Stone. 1990. Privacy in organizations: Theoretical issues, research findings, and protection mechanisms. *Research in Personnel and Human Resources Management*. K. M. Rowland and G. Ferris. JAI Press. **8:** 349-411.

Sundararajan, A. 2004. Managing Digital Piracy: Pricing and Protection. *Information Systems Research* **15**(3)287-308.

Sundararajan, A. 2004. Nonlinear Pricing of Information Goods. *Management Science* **50**(12)1660-1673.

Tsai, J. Y., S. Egelman, L. Cranor and A. Acquisti. forthcoming 2011. The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. *Information Systems Research*.

Turner, N. and L. Wolfson. 2010. *Microsoft to Terminate Bing Cash-Back Program July 30 (Update3). Businessweek* Bloomberg.

Volokh, E. 2000. Personalization and privacy. *Communications of The ACM* **43**(8)84-88.

Yang, H. and L. Ye. 2008. Nonlinear pricing, market coverage, and competition. *Theoretical Economics* **3**(1)123-153.

Appendix A: Figures

Figure 1. Optimal menus under different levels of technology efficiency

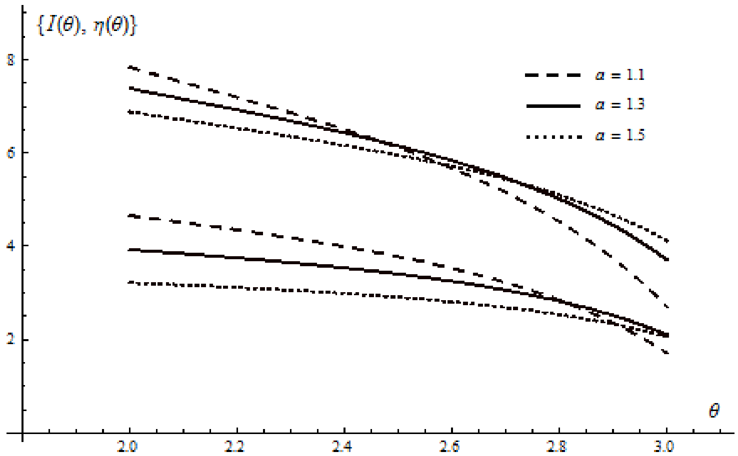


Figure 2. Optimal menu under minimum privacy-preservation policy versus benchmark solution

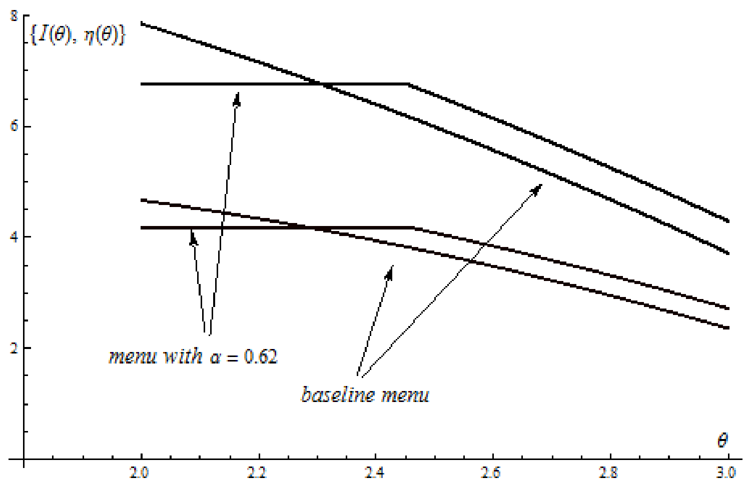


Figure 3. Equilibrium preservation ratios under minimum privacy-preservation policy and in the benchmark case

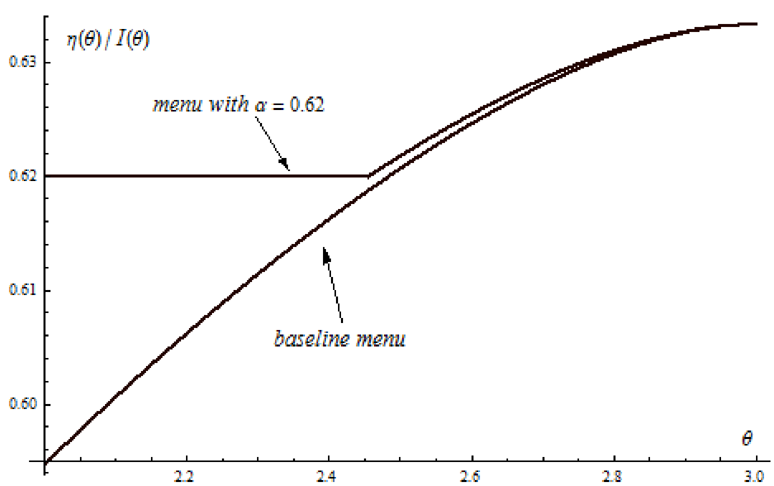


Figure 4. Social welfare under different levels of privacy-preservation regulation

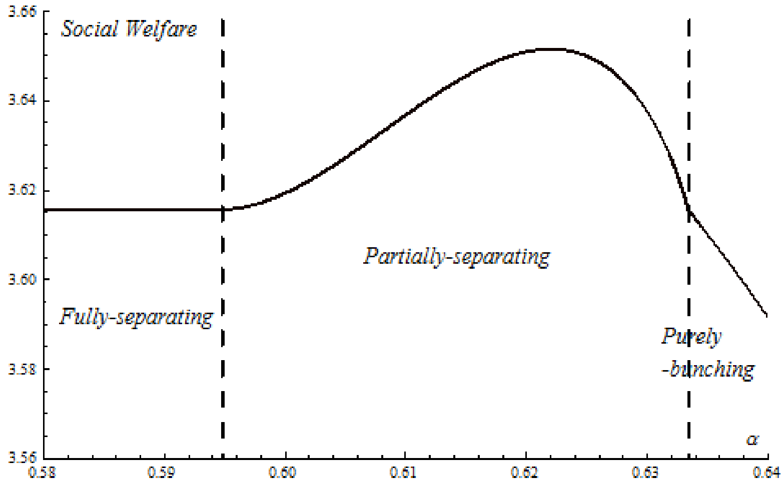
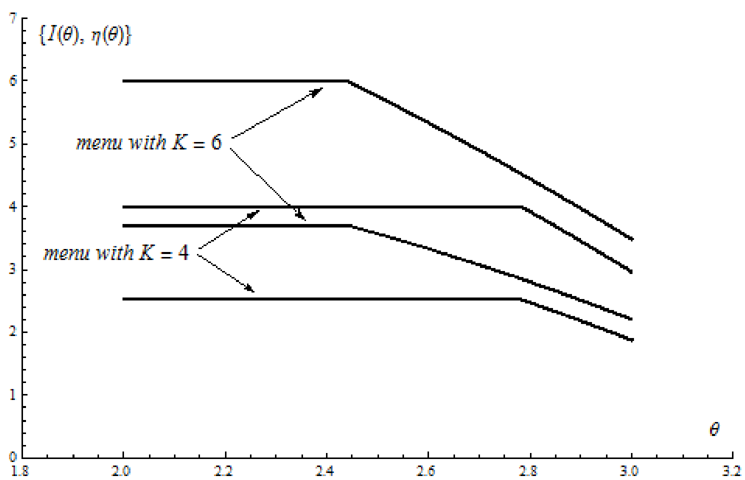


Figure 5. Optimal menus under different levels of information boundary



Appendix B: Proofs

Properties of log-concave density functions

The function $\ln f$ is concave if and only if $\left(\ln f(\theta)\right)'' = \frac{d}{d\theta}\left(\frac{f'(\theta)}{f(\theta)}\right)$ is non-positive.

The function $\ln F$ is concave if and only if $\left(\ln F\right)'' = \frac{d}{d\theta}\left(\frac{f(\theta)}{F(\theta)}\right) = \frac{f'(\theta)F(\theta) - f(\theta)^2}{F(\theta)^2}$ is non-positive.

If $f(\theta)$ is log-concave, then for all $\theta \in [\underline{\theta}, \overline{\theta}]$

$$\frac{f'(\theta)}{f(\theta)}F(\theta) = \frac{f'(\theta)}{f(\theta)}\int_{\underline{\theta}}^{\theta} f(v)dv \le \int_{\underline{\theta}}^{\theta}\frac{f'(v)}{f(v)}f(v)dv = f(\theta) - f(\underline{\theta}) \le f(\theta)$$

Therefore, $f'(\theta)F(\theta)-f(\theta)^2 \le 0$. It also implies $\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right) \ge 0$.

Proof of Lemma 1: $\theta\eta^*(\theta)=(\theta-a)I^*(\theta)$ follows from the vendor's ability to extract all consumer surplus under complete information. The form of the contract menu is a result of the vendor's objective to maximize a concave function of $I(\theta)$. □

Proof of Lemma 2: For $\underline{\theta}>a$, a piecewise C^1 incentive-compatible allocation $\{I(\theta),\eta(\theta)\}$ satisfies $d\eta/d\theta \le 0$ and $(a-\theta)\frac{dI}{d\theta}+\theta\frac{d\eta}{d\theta}=0$ a.e. on $[\underline{\theta},\bar{\theta}]$.

For a direct mechanism of $\{I(\theta),\eta(\theta)\}_{\theta\in[\underline{\theta},\bar{\theta}]}$, truth-telling is the best response from consumers' perspective. This is captured by the IC in equation (3.7), which implies that the following first-order condition is satisfied:

$$a\dot{I}(\tilde{\theta})-\theta\left(\dot{I}(\tilde{\theta})-\dot{\eta}(\tilde{\theta})\right)|_{\tilde{\theta}=\theta}=0$$

or

$$a\dot{I}(\theta)-\theta\left(\dot{I}(\theta)-\dot{\eta}(\theta)\right)=0. \tag{8}$$

Equation (8) holds for all $\theta\in[\underline{\theta},\bar{\theta}]$ since θ is unknown to the principal.

Further, it is also necessary to satisfy the local second-order condition,

$$a\ddot{I}(\tilde{\theta})-\theta\left(\ddot{I}(\tilde{\theta})-\ddot{\eta}(\tilde{\theta})\right)|_{\tilde{\theta}=\theta}\le 0$$

or

$$a\ddot{I}(\theta)-\theta\left(\ddot{I}(\theta)-\ddot{\eta}(\theta)\right)\le 0. \tag{9}$$

By differentiating equation (8), equation (9) can be simplified as

$$\dot{I}(\theta)-\dot{\eta}(\theta)=\dot{i}(\theta)\le 0. \tag{10}$$

Equations (8) and (10) constitute the local incentive constraints, which ensure that the agent has no incentive to lie locally. Based on the single-crossing property, local incentive constraints also imply the global constraints. □

The principal's control problem

Let $\dot{\eta}(\theta)=u$ be the control variable and thus $\dot{I}(\theta)=\frac{\theta}{\theta-a}u$ from lemma 2. The principal's objective represented in equation (3.7) can be transformed into an equivalent problem in the optimal control framework.

$$\int_{\underline{\theta}}^{\bar{\theta}}\left\{b\left(I(\theta)-\eta(\theta)\right)-\frac{1}{2}\left(I(\theta)-\eta(\theta)\right)^2-I(\theta)\right\}f(\theta)d\theta$$

$$s.t. \qquad \dot{\eta}(\theta)=u\le 0 \qquad \dot{I}(\theta)=\frac{\theta}{\theta-a}u \tag{11}$$

$$aI(\theta)-\theta\left(I(\theta)-\eta(\theta)\right)\ge 0 \qquad I(\theta)\ge\eta(\theta) \qquad \eta(\theta)\ge 0$$

We restrict attention to piecewise continuous controls, u, and therefore piecewise smooth schedules. The corresponding Hamiltonian-Lagrangian is defined:

$$L\left(I,\eta,u,p_1,p_2,\lambda,\theta\right) \equiv H\left(I,\eta,u,p_1,p_2,\theta\right) + \lambda_1\left(I(\theta)-\eta(\theta)\right) - \lambda_3\left(\frac{\theta}{\theta-a}u - u\right) - \lambda_4 u\,, \tag{12}$$

where $H\left(I,\eta,u,p_1,p_2,\theta\right) \equiv \left\{b\left(I(\theta)-\eta(\theta)\right) - \frac{1}{2}\left(I(\theta)-\eta(\theta)\right)^2 - I(\theta)\right\} f(\theta) + p_1 u + p_2\frac{\theta}{\theta-a}u$.

In this equation, $I(\theta)$ and $\eta(\theta)$ are state variables, $p_1(\theta)$ and $p_2(\theta)$ are two co-state variables having one-sided limits everywhere, and $\lambda(\theta) = \left(\lambda_1(\theta), \lambda_2(\theta), \lambda_3(\theta), \lambda_4(\theta)\right)$ is a non-decreasing vector function, among which $\lambda_1(\theta)$, $\lambda_3(\theta)$, and $\lambda_4(\theta)$ are associated with the three pure-state constraints in (11), while $\lambda_2(\theta)$ applies to our investigation of the regulatory intervention. Following Seierstad and Sydsæter (1987, Theorems 2 and 3, Chapter 5), we construct the necessary and sufficient conditions for an optimal schedule:

$$u^* = \arg\max_u H\left(I^*,\eta^*,u,p_1,p_2,\theta\right) \text{ for all } u \le 0 \text{ and for a.e. } \theta \in (\underline{\theta},\overline{\theta})\,; \tag{13}$$

$\lambda_j(\theta)$ is constant on any interval where its corresponding constraint is relaxed; and it is continuous at all $\theta \in (\underline{\theta},\overline{\theta})$ where its corresponding constraint binds. Further, u^* is discontinuous (Take $\lambda_1(\theta)$ as an example, this statement expresses the fact that it is constant on any interval where $aI(\theta) - \theta\left(I(\theta)-\eta(\theta)\right) > 0$ at all $\theta \in (\underline{\theta},\overline{\theta})$ at which $aI(\theta) - \theta\left(I(\theta)-\eta(\theta)\right) = 0$ and u^* is discontinuous.)

Define $p_1^* = p_1 + \theta\lambda_1 - \lambda_3 + \lambda_4$, and $p_2^* = p_2 + \left(a-\theta\right)\lambda_1 + \lambda_3$ where p_1^* and p_2^* are strictly continuous, and have derivatives a.e. given by

$$\dot{p}_1^* = -\frac{\partial L}{\partial \eta} = \lambda_1(\theta) - \left(I(\theta)-\eta(\theta)-b\right) f(\theta) \text{ and } \dot{p}_2^* = -\frac{\partial L}{\partial I} = \left(I(\theta)-\eta(\theta)-\left(b-1\right)\right) f(\theta) - \lambda_1(\theta)\,;$$

Finally, $p_1(\overline{\theta}) = p_2(\overline{\theta}) = p_1(\underline{\theta}) = p_2(\underline{\theta}) = 0$ and $\lambda_j(\underline{\theta}) = 0$ (the transversality conditions). By definitions of $p_1^*(\theta)$ and $p_2^*(\theta)$, we have $p_1^*(\underline{\theta}) = p_2^*(\underline{\theta}) = 0$. When combined with the differential equations of the two co-state variables and the continuity of $p_1^*(\theta)$ and $p_2^*(\theta)$, this implies that

$$\begin{cases} p_1^*(\theta) = \int_{\underline{\theta}}^{\theta} \left\{\left(b - \left(I(v)-\eta(v)\right)\right) f(v) + \lambda_1(v)\right\} dv \\ p_2^*(\theta) = \int_{\underline{\theta}}^{\theta} \left\{\left(I(v)-\eta(v)-b+1\right) f(v) - \lambda_1(v)\right\} dv \end{cases}.$$

accordingly,

$$\begin{cases} p_1(\theta) = \int_{\underline{\theta}}^{\theta} \left\{\left(b - \left(I(v)-\eta(v)\right)\right) f(v) + \lambda_1(v)\right\} dv - \theta\lambda_1(\theta) - \lambda_4(\theta) \\ p_2(\theta) = \int_{\underline{\theta}}^{\theta} \left\{\left(I(v)-\eta(v)-b+1\right) f(v) - \lambda_1(v)\right\} dv + \left(\theta-a\right)\lambda_1(\theta) \end{cases} \tag{14}$$

From equation (13), separation (i.e., $u^* < 0$) is only optimal on the interval where $\left(\theta-a\right)p_1 + \theta p_2 = 0$; if there exists some interval over which $\left(\theta-a\right)p_1 + \theta p_2 > 0$, it is optimal for the firm to bunch $\left\{I(\theta),\eta(\theta)\right\}$ on this interval (i.e., when $u^* = 0$).

Lemma A1.

1) For any piecewise C^1 incentive-compatible allocation $\left\{I(\theta),\eta(\theta)\right\}$, $I(\theta) \ge \eta(\theta)$ is irrelevant to the principal's objective if $\eta(\theta) > 0$ a.e. on $[\underline{\theta},\overline{\theta}]$.

2) On optimality, $aI(\theta)-\theta\big(I(\theta)-\eta(\theta)\big)\geq 0$ binds for some θ.

Proof : $\frac{d}{d\theta}\Big[aI(\theta)-\theta\big(I(\theta)-\eta(\theta)\big)\Big]=(a-\theta)\frac{dI}{d\theta}+\theta\frac{d\eta}{d\theta}-\big(I(\theta)-\eta(\theta)\big)=\eta(\theta)-I(\theta)\leq 0$.

The second equality follows from Lemma 2, and the first argument follows from $I(\theta)\geq\eta(\theta)$. It implies that $aI(\theta)-\theta\big(I(\theta)-\eta(\theta)\big)$ is non-increasing in θ on $[\underline{\theta},\overline{\theta}]$.

First, we show that $aI(\theta)-\theta\big(I(\theta)-\eta(\theta)\big)\geq 0$ and $I(\theta)\geq\eta(\theta)$ are exclusive for $I(\theta)>0$ on $[\underline{\theta},\overline{\theta}]$ i.e. if one constraint is active for some θ, the other cannot be active for any $\theta\in[\underline{\theta},\overline{\theta}]$.

To prove this, consider a case where $I(\theta)\geq\eta(\theta)$ binds on some intervals. Lemma 2 implies that $dI/d\theta\leq d\eta/d\theta\leq 0$: once $I(\theta)=\eta(\theta)>0$ at some $\tilde{\theta}$, it continues to bind on the remaining interval $(\tilde{\theta},\overline{\theta}]$. Hence $aI(\theta)-\theta\big(I(\theta)-\eta(\theta)\big)=aI(\theta)>0$ on $(\tilde{\theta},\overline{\theta}]$.

Since $aI(\theta)-\theta\big(I(\theta)-\eta(\theta)\big)$ is non-increasing in θ. $aI(\theta)-\theta\big(I(\theta)-\eta(\theta)\big)>0$ also holds for $\theta\in[\underline{\theta},\tilde{\theta})$. Hence, the participation constraint is inactive on $[\underline{\theta},\overline{\theta}]$. The reverse can be shown in a similar fashion.

Second, we show that vendor prefers binding the participation constraint to $I(\theta)\geq\eta(\theta)$. Suppose on the contrary, the vendor binds $I(\theta)\geq\eta(\theta)$ at some θ. According to the above discussion, $aI(\theta)-\theta\big(I(\theta)-\eta(\theta)\big)>0$ on $[\underline{\theta},\overline{\theta}]$ implies $\lambda_1(\theta)=0$ on $[\underline{\theta},\overline{\theta}]$. Since the transversality conditions and (14) require that $a\lambda_1(\overline{\theta})+\lambda_4(\overline{\theta})=1>0$, it implies $\eta(\theta)\geq 0$ should bind for some θ. A contradiction is derived. □

Since the non-excessive preservation constraint to the vendor's optimal solution is irrelevant, it has been taken out of consideration from our analysis.

Proof of Proposition 1:

From equation (7), $(\theta-a)\,p_1(\theta)+\theta p_2(\theta)=a\int_{\underline{\theta}}^{\theta}\left\{I(v)-\eta(v)-\left(b-\frac{1}{a}\left(v+\frac{F(v)}{f(v)}\right)\right)\right\}f(v)dv$.

We now show that the optimal menu satisfies $(\theta-a)\,p_1(\theta)+\theta p_2(\theta)=0$ on $[\underline{\theta},\overline{\theta}]$. Suppose on the contrary, there exists a point $\tilde{\theta}$ at which $\big[(\theta-a)\,p_1(\theta)+\theta p_2(\theta)\big]'\,|_{\theta=\tilde{\theta}^+}>0$. This implies

$I(\tilde{\theta})-\eta(\tilde{\theta})-\left(b-\frac{1}{a}\left(\tilde{\theta}+\frac{F(\tilde{\theta})}{f(\tilde{\theta})}\right)\right)>0$.

By continuity, for an arbitrarily small δ, $(\theta-a)\,p_1(\theta)+\theta p_2(\theta)>0$ on $(\tilde{\theta},\tilde{\theta}+\delta)$.

Based on equation (6) in the Online Appendix, we have $u^*=0$ on $(\tilde{\theta},\tilde{\theta}+\delta)$ (i.e., $I(\theta)$ and $\eta(\theta)$ keep constant on $(\tilde{\theta},\tilde{\theta}+\delta)$). Since $\frac{F(\theta)}{f(\theta)}$ is non-decreasing in θ, $I(\tilde{\theta})-\eta(\tilde{\theta})>b-\frac{1}{a}\left(\theta+\frac{F(\theta)}{f(\theta)}\right)$ holds on $(\tilde{\theta},\overline{\theta}]$, as well as $(\theta-a)\,p_1(\theta)+\theta p_2(\theta)>0$. Together they imply bunching till $\overline{\theta}$. As a consequence,

$$p_1(\overline{\theta})=\int_{\underline{\theta}}^{\overline{\theta}}\big\{b-\big(I(v)-\eta(v)\big)\big\}f(v)dv-\overline{\theta}q_1(\overline{\theta})=\int_{\tilde{\theta}}^{\overline{\theta}}\left\{b-\frac{1}{a}\left(v+\frac{F(v)}{f(v)}\right)-\big(I(\tilde{\theta})-\eta(\tilde{\theta})\big)\right\}f(v)dv<0$$

This results in a contradiction to the transversality conditions, which require $p_1(\overline{\theta})=0$ on optimality.

The optimal menu specified in Proposition 1 results directly from the intermediate result that $(\theta - a)\, p_1(\theta) + \theta p_2(\theta) = 0$ on $[\underline{\theta}, \overline{\theta}]$ and the fact the participation constraint has to bind at $\overline{\theta}$ from Lemma A1. It is fully-revealing with $u(\theta) = -\frac{\theta - a}{a^2}\left(\theta + \frac{F(\theta)}{f(\theta)}\right)'$. □

Proof of Proposition 2 (refer to "free terminal problem" in Seierstad and Sydsæter (1987)):
Given an arbitrary $\theta_T \in [\underline{\theta}, \overline{\theta}]$ characterizing the market coverage of the optimal menu. There exists an admissible pair $(I_T(\theta), \eta_T(\theta), u_T(\theta))$ defined on $[\underline{\theta}, \theta_T]$ with associated multipliers $p_T(\theta)$ and $q_T(\theta)$ satisfying all the sufficient conditions for an optimal control problem.

$$\begin{cases} p_{1T}(\theta) = \int_{\underline{\theta}}^{\theta} \left\{ \left(b - \left(I_T(v) - \eta_T(v)\right)\right) f(v) + \lambda_{1T}(v) \right\} dv - \theta \lambda_{1T}(\theta) - \lambda_{4T}(\theta) \\ p_{2T}(\theta) = \int_{\underline{\theta}}^{\theta} \left\{ \left(I_T(v) - \eta_T(v) - b + 1\right) f(v) - \lambda_{1T}(v) \right\} dv + (\theta - a)\, \lambda_{1T}(\theta) \end{cases} \text{ and } p_{1T}(\theta_T) = p_{2T}(\theta_T) = 0 .$$

Define

$H_T(\theta_T^-) = \left\{ b\left(I_T(\theta_T) - \eta_T(\theta_T)\right) - \frac{1}{2}\left(I_T(\theta_T) - \eta_T(\theta_T)\right)^2 - I_T(\theta_T) \right\} f(\theta_T) + \left(p_{1T}(\theta_T^-) + p_{2T}(\theta_T^-) \frac{\theta_T}{\theta_T - a} \right) u_T(\theta_T^-)$ Since $H_T(\theta_T^-) + \left(q_{1T}(\theta_T) - q_{1T}(\theta_T^-)\right)\left(-\left(I_T(\theta_T) - \eta_T(\theta_T)\right)\right) > 0$, $\theta_T^* = \overline{\theta}$.

If $\overline{\theta} < ab < \overline{\theta} + \frac{F(\overline{\theta})}{f(\overline{\theta})}$, there exists $\breve{\theta}$ satisfying $ab = \breve{\theta} + \frac{F(\breve{\theta})}{f(\breve{\theta})}$. For an arbitrary θ_T satisfying $\theta_T \geq \breve{\theta}$.

On $[\underline{\theta}, \breve{\theta})$, $I(\theta) - \eta(\theta) = b - \frac{1}{a}\left(\theta + \frac{F(\theta)}{f(\theta)}\right)$, whereas on $[\breve{\theta}, \theta_T]$, $I(\theta) = \eta(\theta) = 0$. $aq_{1T}(\theta_T) + q_{4T}(\theta_T) = F(\theta_T)$ $H_T(\theta_T^-) + \left(q_{1T}(\theta_T) - q_{1T}(\theta_T^-)\right)\left(-\left(I_T(\theta_T) - \eta_T(\theta_T)\right)\right) = 0$. For an arbitrary θ_T satisfying $\theta_T < \breve{\theta}$,

$$H_T(\theta_T^-) + \left(q_{1T}(\theta_T) - q_{1T}(\theta_T^-)\right)\left(-\left(I_T(\theta_T) - \eta_T(\theta_T)\right)\right) = \frac{f(\theta_T)}{2}\left(b - \frac{1}{a}\left(\theta_T + \frac{F(\theta_T)}{f(\theta_T)}\right)\right)^2 > 0 .$$

Together, the point that delimits the shutdown portion is $\breve{\theta}$, which is the solution of $ab = \breve{\theta} + \frac{F(\breve{\theta})}{f(\breve{\theta})}$ □

Lemma A2. Define the auxiliary function $P(\theta, \mu) = \theta + \frac{F(\theta) + \mu}{f(\theta)}$ for $\theta \in [\underline{\theta}, \overline{\theta}]$ where μ is an arbitrary constant. The function $P(\theta, \mu)$ is quasi-convex in θ on $[\underline{\theta}, \overline{\theta}]$ if $\mu \geq 0$; and is increasing in θ on $[\underline{\theta}, \overline{\theta}]$ if $-1 \leq \mu < 0$.

Proof:
We only prove the first part of this lemma. The proof of the latter follows a symmetric argument.

To substantiate this claim, consider an arbitrary extremum (providing one exists) $\hat{\theta}$ of $P(\theta, \mu)$. Note that

$$f(\theta)^3 P_{\theta\theta}(\theta,\mu) = f(\theta)^3\left(\frac{2f(\theta)^2-(F(\theta)+\mu)f'(\theta)}{f(\theta)^2}\right)'$$

$$=\frac{\left(4f(\theta)f'(\theta)-(F(\theta)+\mu)f''(\theta)-f'(\theta)f(\theta)\right)f(\theta)^2-2\left(2f(\theta)^2-(F(\theta)+\mu)f'(\theta)\right)f(\theta)f'(\theta)}{f(\theta)}$$

$$=-f(\theta)^2 f'(\theta)+\left(2f'(\theta)^2-f''(\theta)f(\theta)\right)\left(F(\theta)+\mu\right)$$

The condition that defines an extremum is given by $P_\theta(\hat{\theta},\mu)=2-\frac{(F(\hat{\theta})+\mu)f'(\hat{\theta})}{f(\hat{\theta})^2}=0$, which can be the case only when $f'(\hat{\theta})>0$ (since $F(\theta)+\mu\geq 0$ for all $\theta\in[\underline{\theta},\overline{\theta}]$).

Using $P_\theta(\hat{\theta},\mu)=2-\frac{(F(\hat{\theta})+\mu)f'(\hat{\theta})}{f(\hat{\theta})^2}=0$ to simplify the above second-order expression, we have

$$P_{\theta\theta}(\hat{\theta},\mu)=\frac{1}{f(\hat{\theta})f'(\hat{\theta})}\left(f'(\hat{\theta})^2+2\left(f'(\hat{\theta})^2-f''(\hat{\theta})f(\hat{\theta})\right)\right)>0,$$

where the inequality follows from the fact that $\frac{d}{d\theta}\left(\frac{f'(\theta)}{f(\theta)}\right)$ is non-positive and $f'(\theta)>0$.

Therefore, for $\mu\geq 0$, if there exists any extremum of $P(\theta,\mu)$, it must be a minimum. In other words, $P(\theta,\mu)$ is single-peaked. Since $f(\theta)$ is continuously differentiable and positive everywhere on the support, $P_\theta(\theta,\mu)$ is continuous and well-defined. Though an extremum does not necessarily exist, once $P_\theta(\theta,\mu)\geq 0$, it remains so for all higher values of θ.

When $-1\leq\mu<0$, define $\hat{\theta}$ as the point at which $F(\hat{\theta})+\mu=0$. If there exists an extremum in the interval $[0,\hat{\theta})$, $P_\theta(\hat{\theta},\mu)=2-\frac{(F(\hat{\theta})+\mu)f'(\hat{\theta})}{f(\hat{\theta})^2}=0$ requires $f'(\hat{\theta})<0$, implying that

$$P_{\theta\theta}(\hat{\theta},\mu)=\frac{1}{f(\hat{\theta})f'(\hat{\theta})}\left(f'(\hat{\theta})^2+2\left(f'(\hat{\theta})^2-f''(\hat{\theta})f(\hat{\theta})\right)\right)<0.$$

In other words, this extremum must be a maximum. Then at the point $\hat{\theta}$ where $F(\hat{\theta})+\mu=0$, $P_\theta(\hat{\theta},\mu)<0$ by continuity. However, $P_\theta(\hat{\theta},\mu)=2-\frac{(F(\hat{\theta})+\mu)f'(\hat{\theta})}{f(\hat{\theta})^2}>0$. A contradiction is derived, implying that there cannot be an extremum in the interval $[0,\hat{\theta})$.

In the interval $[\hat{\theta},\overline{\theta}]$, $F(\theta)+\mu>0$. $P_\theta(\theta,\mu)>0$ because $P_\theta(\hat{\theta},\mu)>0$ according to the first part of this lemma. □

Minimum Privacy-Preservation Policy

To investigate the effects of a minimum preservation ratio on the optimal schedule, we associate inequality in equation (4.1) with the Lagrange multiplier $\lambda_2(\theta)$. The new Hamiltonian-Lagrangian is given by:

$$L\left(I,\eta,u,p_1,p_2,\lambda,\theta\right)\equiv H+\lambda_1\left(I(\theta)-\eta(\theta)\right)-\lambda_2\left(u-\frac{\alpha\theta}{\theta-a}u\right).$$

The necessary condition is similar to that in the benchmark case, except that $p_1^*=p_1+\theta\lambda_1+\lambda_2$ and $p_2^*=p_2+(a-\theta)\lambda_1-\alpha\lambda_2$ here. Accordingly, the counterpart of equation (7) is

$$\begin{cases} p_1(\theta) = \int_{\underline{\theta}}^{\theta} \left\{ \left(b - \left(I(v) - \eta(v) \right) \right) f(v) + \lambda_1(v) \right\} dv - \theta \lambda_1(\theta) - \lambda_2(\theta) \\ p_2(\theta) = \int_{\underline{\theta}}^{\theta} \left\{ \left(I(v) - \eta(v) - b + 1 \right) f(v) - \lambda_1(v) \right\} dv + \left(\theta - a \right) \lambda_1(\theta) + \alpha \lambda_2(\theta) \end{cases}.$$

To facilitate the analysis of the optimal schedule, we shall first present the following intermediate results:

$$\begin{aligned} &\left(\theta - a \right) p_1(\theta) + \theta p_2(\theta) \\ &= a \int_{\underline{\theta}}^{\theta} \left\{ \left(I(v) - \eta(v) - \left(b - \frac{\theta}{a} \right) \right) f(v) - \lambda_1(v) \right\} dv + \left(a - \theta \left(1 - \alpha \right) \right) \lambda_2(\theta) \end{aligned} \tag{15}$$

Transversality conditions imply

$$1 = a \lambda_1(\bar{\theta}) + \left(1 - \alpha \right) \lambda_2(\bar{\theta}) \tag{16}$$

which suggests that at least one of the participation constraint and the minimum preservation constraint bind on optimality.

It can be verified that $d\left(\frac{\tilde{\eta}(\theta)}{\tilde{I}(\theta)} \right) / d\theta \geq 0$ with the equality holding only at $\bar{\theta}$. Define a function $g(\theta) = \frac{\tilde{\eta}(\theta)}{\tilde{I}(\theta)}$, then we have $g(\underline{\theta}) = \underline{\alpha}$, and $g(\bar{\theta}) = \frac{\bar{\theta} - a}{\bar{\theta}}$.

The followings examine the monotonicity of $\eta(\theta) - \alpha I(\theta)$.

From Lemma 2, $\frac{d}{d\theta}\left(\eta(\theta) - \alpha I(\theta) \right) = \dot{\eta}(\theta) - \alpha \dot{I}(\theta) = u - \frac{\alpha\theta}{\theta - a} u$ with $u \leq 0$, from which we observe that the monotonicity of $\eta(\theta) - \alpha I(\theta)$ depends on the magnitude of $\frac{\alpha\theta}{\theta - a}$ as opposed to 1 . $\frac{\alpha\theta}{\theta - a}$ is decreasing in θ.

Define an increasing function $h(\theta) = \frac{\theta - a}{\theta}$. Let $\bar{\alpha} = h(\bar{\theta}) = \frac{\bar{\theta} - a}{\bar{\theta}}$. $\bar{\alpha} = g(\bar{\theta}) > \underline{\alpha} = g(\underline{\theta})$ is ensured by the fact that $g'(\theta) > 0$. Since $\frac{\bar{\alpha}\bar{\theta}}{\bar{\theta} - a} = 1$, $\eta(\theta) - \alpha I(\theta)$ is non-decreasing on $[\underline{\theta}, \bar{\theta}]$ for any $\alpha \in [\bar{\alpha}, 1)$.

When $\alpha \in (\underline{\alpha}, \bar{\alpha})$, there exists $\tilde{\theta} = h^{-1}(\alpha) = \frac{a}{1 - \alpha} < \bar{\theta}$ such that $\eta(\theta) - \alpha I(\theta)$ is non-decreasing on $[\underline{\theta}, \tilde{\theta}]$, and is non-increasing on $(\tilde{\theta}, \bar{\theta}]$;

A useful result that facilitates our analysis is that $h^{-1}(\alpha) > g^{-1}(\alpha)$ for any $\alpha \in (\underline{\alpha}, \bar{\alpha})$, which follows the fact that $h(\theta)$ is smaller than $g(\theta)$ for any θ less than $\bar{\theta}$. Thereafter, for any $\alpha \in (\underline{\alpha}, \bar{\alpha})$, there always exists an interval $(g^{-1}(\alpha), h^{-1}(\alpha))$ over which the constraint is slack for the benchmark solution.

Proof of Lemma 3:

The proof proceeds in two steps. First, we show that full separation is no longer optimal.

Suppose on the contrary, i.e. full separation is still optimal, we should have $(\theta - a)p_1 + \theta p_2 = 0$ over $[\underline{\theta}, \overline{\theta}]$. Governed by $u < 0$, the participation constraint can only bind at $\overline{\theta}$. For a given $\alpha \in (\underline{\alpha}, \overline{\alpha})$, $\eta(\theta) - \alpha I(\theta) \geq 0$ is slack on $(\underline{\theta}, \overline{\theta}]$; otherwise, the participation constraint is violated in a neighborhood of $\overline{\theta}$, because $\eta(\theta)/I(\theta) = \alpha < \frac{\overline{\theta} - a}{\overline{\theta}}$. So $\lambda_2(\theta)$ is constant on $(\underline{\theta}, \overline{\theta}]$. Using λ_2 to denote the constant value of $\lambda_2(\theta)$ on $(\underline{\theta}, \overline{\theta}]$, we obtain:

$$(\theta - a)\, p_1(\theta) + \theta p_2(\theta) = a\int_{\underline{\theta}}^{\theta} \left(I(v) - \eta(v) - \left(b - \frac{1}{a}\left(v + \frac{F(v) - (1-\alpha)\lambda_2}{f(v)} \right) \right) \right) f(v)dv + \left(a - \underline{\theta}(1-\alpha) \right)\lambda_2 = 0 \quad \forall \theta \in (\underline{\theta}, \overline{\theta}]$$

which implies $I^*(\theta) - \eta^*(\theta) = b - \frac{1}{a}\left(\theta + \frac{F(\theta) - (1-\alpha)\lambda_2}{f(\theta)} \right)$ and $\lambda_2 = 0$ (because of $\frac{a}{1-\alpha} > \underline{\theta}$ for $\alpha \in (\underline{\alpha}, \overline{\alpha})$) over $[\underline{\theta}, \overline{\theta}]$.

Hence, $u^*(\theta) = -\frac{\theta - a}{a^2}\left(\theta + \frac{F(\theta)}{f(\theta)} \right)' < 0$, the same as the benchmark case; so is $\left\{ I^*(\theta), \eta^*(\theta) \right\}$ (because the participation constraint also binds at $\overline{\theta}$ as it does in the benchmark case). A contradiction is derived. In addition, for $\alpha \in [\overline{\alpha}, 1)$, only purely bunching menu can be sustained.

Second, we show that the principal's optimal menu will exhibit bunching at most on an interval at the bottom of the type space.

Suppose that $(\theta', \theta'') \subset [\underline{\theta}, \overline{\theta}]$ is the first interval from below over which the optimal allocation is full separation (define θ' as $\theta' = \inf\left\{ t : u^*(\theta) < 0 \text{ for all } t < \theta < \theta'' \right\}$ and θ'' as $\theta'' = \sup\left\{ t : u^*(\theta) < 0 \text{ for all } \theta' < \theta < t \right\}$).

Because u^* is discontinuous at θ' and θ'', $\lambda_2(\theta)$ is continuous at the two points, so is $(\theta - a)\, p_1(\theta) + \theta p_2(\theta)$ from equation (15). Hence we obtain $(\theta - a)\, p_1(\theta) + \theta p_2(\theta) = 0$ on $[\theta', \theta'']$. $\eta(\theta) \geq \alpha I(\theta)$ is slack over (θ', θ''), implying that $\lambda_2(\theta)$ is constant over (θ', θ''). We use λ_2 to denote the constant value of $\lambda_2(\theta)$ on (θ', θ''), and show that θ'' should coincide with $\overline{\theta}$.

Since $(\theta - a)\, p_1(\theta) + \theta p_2(\theta) = 0$ over (θ', θ''),

$$I(\theta) - \eta(\theta) = b - \frac{1}{a}\left(\theta + \frac{F(\theta) - (1-\alpha)\lambda_2}{f(\theta)} \right) \quad \forall \theta \in (\theta', \theta'').$$

It is equivalent to have $u^*(\theta) = -\frac{\theta - a}{a^2}\frac{\partial}{\partial \theta}\left(\theta + \frac{F(\theta) - (1-\alpha)\lambda_2}{f(\theta)} \right)$ on (θ', θ''). Lemma A2 shows that given $0 < (1-\alpha)\lambda_2 \leq 1$ (by equation (9)), $\frac{\partial}{\partial \theta}\left(\theta + \frac{F(\theta) - (1-\alpha)\lambda_2}{f(\theta)} \right) > 0$ holds on the entire type space.

Next, we show that if θ'' is different from $\overline{\theta}$, $\eta(\theta) \geq \alpha I(\theta)$ cannot bind at θ''. Suppose otherwise that $\eta(\theta) = \alpha I(\theta)$ at θ'', θ'' should rest on the non-increasing region of $\eta(\theta) - \alpha I(\theta)$, implying $\theta'' > \frac{a}{1-\alpha}$. Because $\alpha \in (\underline{\alpha}, \overline{\alpha})$, $aI(\theta) - \theta\left(I(\theta) - \eta(\theta) \right) = (1-\alpha)\left(\frac{a}{(1-\alpha)} - \theta \right) I(\theta) < 0$ on $[\theta'', \overline{\theta}]$, i.e. the participation constraint is violated.

Therefore, $\eta(\theta) \geq \alpha I(\theta)$ should remain slack on $[\theta'', \bar{\theta}]$, and $\lambda_2(\theta)$ is constant over $(\theta', \bar{\theta}]$. Bunching till $\bar{\theta}$ is sustained by $(\theta - a)p_1(\theta) + \theta p_2(\theta) > 0$ on $[\theta'', \bar{\theta}]$. Then it is obtained:

$$p_1(\bar{\theta}) = \int_{\underline{\theta}}^{\bar{\theta}} \left(b - \left(I(v) - \eta(v)\right)\right) f(v)dv - \bar{\theta}\lambda_1(\bar{\theta}) - \lambda_2(\bar{\theta})$$

$$= \int_{\underline{\theta}}^{\bar{\theta}} \left(b - \left(I(v) - \eta(v)\right)\right) f(v)dv - \frac{\bar{\theta} - \bar{\theta}(1-\alpha)\lambda_2 + a\lambda_2}{a}$$

$$= \int_{\theta''}^{\bar{\theta}} \left(b - \frac{1}{a}\left(v + \frac{F(v) - (1-\alpha)\lambda_2}{f(v)}\right) - \left(I(v) - \eta(v)\right)\right) f(v)dv$$

$$= \frac{1}{a}\int_{\theta''}^{\bar{\theta}} \left(\left(\theta'' + \frac{F(\theta'') - (1-\alpha)\lambda_2}{f(\theta'')}\right) - \left(v + \frac{F(v) - (1-\alpha)\lambda_2}{f(v)}\right)\right) f(v)dv < 0$$

The first equality follows equation(16); the second equality follows $(\theta'' - a)p_1(\theta'') + \theta'' p_2(\theta'') = 0$; and the last inequality follows Lemma A2. Hence the transversality conditions are violated.

Combining these two steps, we complete the proof for Lemma 3. □

Proof of Proposition 3:

From Lemma 3, the principal bunches the low-end of the type space $[\underline{\theta}, \theta']$. θ' here plays a similar role that $\hat{\theta}$ does in the main text. Equation (4.2) is derived from $(\theta' - a)p_1(\theta') + \theta' p_2(\theta') = 0$, and equation (4.3) results from the binding of the participation constraint at $\bar{\theta}$ and the binding of the minimum preservation constraint on $[\underline{\theta}, \theta']$.

According to equation (4.2), $(1-\alpha)\lambda = \dfrac{F(\hat{\theta})^2}{f(\hat{\theta})\left(\dfrac{a}{(1-\alpha)} - \hat{\theta} + \dfrac{F(\hat{\theta})}{f(\hat{\theta})}\right)}$.

$F(\hat{\theta}) - (1-\alpha)\lambda = F(\hat{\theta})\left(\dfrac{\dfrac{a}{(1-\alpha)} - \hat{\theta}}{\dfrac{a}{(1-\alpha)} - \hat{\theta} + \dfrac{F(\hat{\theta})}{f(\hat{\theta})}}\right) \geq 0$ follows $\dfrac{a}{(1-\alpha)} - \hat{\theta} \geq 0$ by equation (4.3).□

Proof of Corollary 1:

1. According to the implicit function theorem, the following two equations are derived from equations (4.2) and (4.3):

$$\left(\hat{\theta} - \frac{F(\hat{\theta})}{f(\hat{\theta})}\right)\lambda + \left[(1-\alpha)\frac{F(\hat{\theta})}{f(\hat{\theta})} + \left(a - \hat{\theta}(1-\alpha)\right)\right]\frac{\partial \lambda}{\partial \alpha} - \left[\left(\hat{\theta} + \frac{F(\hat{\theta}) - (1-\alpha)\lambda}{f(\theta')}\right)' F(\hat{\theta})\right]\frac{\partial \hat{\theta}}{\partial \alpha} = 0$$

$$\left[\left(\frac{a}{(1-\alpha)} - \hat{\theta}\right)\left(\hat{\theta} + \frac{F(\hat{\theta}) - (1-\alpha)\lambda}{f(\hat{\theta})}\right)'\right]\frac{\partial \hat{\theta}}{\partial \alpha} + \frac{(1-\alpha)}{f(\hat{\theta})}\left[\left(\hat{\theta} - \frac{a}{(1-\alpha)}\right) + \int_{\theta'}^{\bar{\theta}} \frac{f(\hat{\theta})}{f(v)}dv\right]\frac{\partial \lambda}{\partial \alpha}$$

$$+ \frac{\lambda}{f(\hat{\theta})}\left(\frac{a}{(1-\alpha)} - \hat{\theta}\right) - \int_{\hat{\theta}}^{\bar{\theta}} \frac{\lambda}{f(v)}dv - \frac{a^2}{(1-\alpha)^2}\left(b - \frac{1}{a}\left(\hat{\theta} + \frac{F(\hat{\theta}) - (1-\alpha)\lambda}{f(\hat{\theta})}\right)\right) = 0$$

which can be reduced to

$$\left[\frac{(1-\alpha)}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)'\left(\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)^2+\int_{\theta'}^{\bar{\theta}}\frac{F(\hat{\theta})}{f(v)}dv\right)\right]\frac{\partial\hat{\theta}}{\partial\alpha}$$

$$=\int_{\hat{\theta}}^{\bar{\theta}}\frac{\lambda}{f(v)}dv+\frac{a}{(1-\alpha)}\left(b-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})}{f(\hat{\theta})}\right)\right)\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)+\frac{a}{(1-\alpha)}\left(b-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)\right)\frac{F(\hat{\theta})}{f(\hat{\theta})}$$

Hence we can conclude that $\frac{\partial\hat{\theta}}{\partial\alpha}>0$.

It can also be verified that $\hat{\theta}^{-1}(\underline{\theta})=\underline{\alpha}$ and $\hat{\theta}^{-1}(\bar{\theta})=\frac{\bar{\theta}-a}{\bar{\theta}}=\bar{\alpha}$ from equations (4.2) and (4.3).

When $\alpha>\bar{\alpha}$, the whole type space would be bunched with $\alpha I^P=\eta^P$. The participation constraint $aI^P-\theta\left(I^P-\eta^P\right)=(1-\alpha)\left(\frac{a}{1-\alpha}-\theta\right)I^P>0$ on the entire type space.

The principal's objective then becomes $\int_{\underline{\theta}}^{\bar{\theta}}\left(b(1-\alpha)I^P-\frac{1}{2}\left((1-\alpha)I^P\right)^2-I^P\right)f(\theta)d\theta$. The first-order condition gives the solution presented in the second part of this corollary. □

Proof of Proposition 4:

The proof of this proposition consists of several arguments. First, we show that $SW'(\alpha)|_{\alpha=\underline{\alpha}}=0$ and $SW''(\alpha)|_{\alpha=\underline{\alpha}}>0$. They imply that an effective intervention on the preservation ratio is a social welfare-improving strategy; for $\alpha\in[\underline{\alpha},\bar{\alpha})$, the first-order derivative of the social planner's objective w.r.t. α:

$$SW'(\alpha)=\left[\left(-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)'\right)\int_{\underline{\theta}}^{\hat{\theta}}\left(\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)-\frac{\theta}{a}\right)f(\theta)d\theta\right]\frac{d\hat{\theta}(\alpha)}{d\alpha}$$

$$+\left[\frac{1}{a}\frac{(1-\alpha)}{f(\hat{\theta})}\int_{\underline{\theta}}^{\hat{\theta}}\left(\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)-\frac{\theta}{a}\right)f(\theta)d\theta+\frac{(1-\alpha)}{a}\int_{\hat{\theta}}^{\bar{\theta}}\frac{1}{a}\left(\frac{F(\theta)-(1-\alpha)\lambda}{f(\theta)}\right)d\theta\right]\frac{d\lambda(\alpha)}{d\alpha}$$

$$-\frac{1}{a}\frac{\lambda}{f(\hat{\theta})}\int_{\underline{\theta}}^{\hat{\theta}}\left(\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)-\frac{\theta}{a}\right)f(\theta)d\theta-\frac{\lambda}{a}\int_{\hat{\theta}}^{\bar{\theta}}\frac{1}{a}\left(\frac{F(\theta)-(1-\alpha)\lambda}{f(\theta)}\right)d\theta$$

Substitute $\hat{\theta}(\underline{\alpha})=\underline{\theta}$ and $\lambda(\underline{\alpha})=0$ into this expression, we obtain $SW'(\underline{\alpha})=0$.

After substituting $\frac{d\hat{\theta}(\alpha)}{d\alpha}$ and $\frac{d\lambda(\alpha)}{d\alpha}$ with their respective values, the first-order derivatives can be simplified as

$$\frac{\frac{a}{(1-\alpha)}}{a^2\left[\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)^2+\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\hat{\theta})}{f(v)}dv\right]}\left\{\begin{array}{l}-\left[\int_{\underline{\theta}}^{\hat{\theta}}\left(\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)-\theta\right)f(\theta)d\theta\right]\left\{\frac{a}{(1-\alpha)}\left(b-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)\right)\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)+\lambda\int_{\hat{\theta}}^{\bar{\theta}}\frac{1}{f(v)}dv\right\}\\ +\left\{F(\hat{\theta})\frac{a}{(1-\alpha)}\left(b-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)\right)-\lambda\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)\right\}\int_{\hat{\theta}}^{\bar{\theta}}\left(\frac{F(\theta)-(1-\alpha)\lambda}{f(\theta)}\right)d\theta\end{array}\right\}$$ with

the term in the outmost bracket equals to 0 at $\alpha=\underline{\alpha}$. Further, differentiating $SW'(\alpha)$ w.r.t. α and evaluating its value at $\alpha=\underline{\alpha}$ yield $SW''(\underline{\alpha})=\dfrac{a^2 f(\underline{\theta})\left(b-\frac{\underline{\theta}}{a}\right)^2\int_{\underline{\theta}}^{\bar{\theta}}\frac{F(\theta)}{f(\theta)}d\theta}{2(1-\alpha)^4\left(\frac{a}{(1-\alpha)}-\underline{\theta}\right)^3}>0$.

Second, we show that $SW'(\alpha)|_{\alpha=\bar{\alpha}^+}<0$ and $SW''(\alpha)<0$ on $\alpha\in[\bar{\alpha},1)$; that is, the social planner prefers the corner solution among all purely-bunching menus; the result can be easily derived by substituting $I^B=\dfrac{1}{(1-\alpha)}\left(b-\dfrac{1}{(1-\alpha)}\right)$ into the social planner's objective for $\alpha\in[\bar{\alpha},1)$, and differentiating it w.r.t. α.

Third, we show that $SW'(\alpha)|_{\alpha=\bar{\alpha}^-}<0$; that is, the social planner's prefers a partially-separating market outcome. For $\alpha\in[\underline{\alpha},\bar{\alpha})$,

$$SW'(\alpha)$$
$$=\frac{\frac{a}{(1-\alpha)}}{a^2\left[\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)^2+\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\hat{\theta})}{f(v)}dv\right]}\left\{\begin{array}{l}\frac{a}{(1-\alpha)}\left(b-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)\right)\left\{F(\hat{\theta})\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\theta)}{f(\theta)}d\theta-\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)\int_{\underline{\theta}}^{\hat{\theta}}F(\theta)d\theta\right\}\\ -\lambda\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\theta)}{f(\theta)}d\theta-\lambda\int_{\hat{\theta}}^{\bar{\theta}}\frac{1}{f(\theta)}d\theta\left(\int_{\underline{\theta}}^{\hat{\theta}}F(\theta)d\theta\right)\end{array}\right\}-\frac{\lambda}{(1-\alpha)}\left(b-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)\right)$$

When $\alpha\to\bar{\alpha}=\dfrac{\bar{\theta}-a}{\bar{\theta}}$ from below, $\left(\dfrac{a}{(1-\alpha)}-\hat{\theta}\right)^2+\displaystyle\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\hat{\theta})}{f(v)}dv\to 0$ as well as

$$\left\{\begin{array}{l}\frac{a}{(1-\alpha)}\left(b-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)\right)\left\{F(\hat{\theta})\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\theta)}{f(\theta)}d\theta-\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)\int_{\underline{\theta}}^{\hat{\theta}}F(\theta)d\theta\right\}\\ -\lambda\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\theta)}{f(\theta)}d\theta-\lambda\int_{\hat{\theta}}^{\bar{\theta}}\frac{1}{f(\theta)}d\theta\left(\int_{\underline{\theta}}^{\hat{\theta}}F(\theta)d\theta\right)\end{array}\right\}\to 0.$$

Moreover, $\dfrac{d\hat{\theta}}{d\alpha}\to\infty$. According to L'Hôpital's rule,

$$\lim_{\alpha\to\bar{\alpha}^-}SW'(\alpha)=\frac{\frac{d}{d\alpha}\left\{\begin{array}{l}\left(b-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)\right)\left\{F(\hat{\theta})\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\theta)}{f(\theta)}d\theta-\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)\int_{\underline{\theta}}^{\hat{\theta}}F(\theta)d\theta\right\}\\ -\frac{(1-\alpha)}{a}\lambda\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\theta)}{f(\theta)}d\theta-\frac{(1-\alpha)}{a}\lambda\int_{\hat{\theta}}^{\bar{\theta}}\frac{1}{f(\theta)}d\theta\left(\int_{\underline{\theta}}^{\hat{\theta}}F(\theta)d\theta\right)\end{array}\right\}}{\frac{d}{d\alpha}\left\{(1-\alpha)^2\left[\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)^2+\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\hat{\theta})}{f(v)}dv\right]\right\}}-\frac{\lambda}{(1-\alpha)}\left(b-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)\right)$$

The nominator of the first term equals as $\alpha\to\bar{\alpha}^-$:

$$\frac{d}{d\alpha}\left\{\begin{array}{l}\left[\left(b-\frac{1}{a}\left(\hat{\theta}+\frac{F(\hat{\theta})-(1-\alpha)\lambda}{f(\hat{\theta})}\right)\right)\left\{F(\hat{\theta})\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\theta)}{f(\theta)}d\theta-\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)\int_{\underline{\theta}}^{\hat{\theta}}F(\theta)d\theta\right\}\right] \\ -\frac{(1-\alpha)}{a}\lambda\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\theta)}{f(\theta)}d\theta-\frac{(1-\alpha)}{a}\lambda\int_{\hat{\theta}}^{\bar{\theta}}\frac{1}{f(\theta)}d\theta\left(\int_{\underline{\theta}}^{\hat{\theta}}F(\theta)d\theta\right)\end{array}\right\}$$

$$=\frac{d\hat{\theta}(\alpha)}{d\alpha}+0\frac{d\lambda(\alpha)}{d\alpha}-\frac{a}{(1-\alpha)^2}\left(b-\frac{\bar{\theta}}{a}\right)\int_{\underline{\theta}}^{\bar{\theta}}F(\theta)d\theta$$

The denominator of the first term equals as $\alpha\to\bar{\alpha}^-$:

$$(1-\alpha)^2\left[-2\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)+\int_{\hat{\theta}}^{\bar{\theta}}\frac{f(\hat{\theta})}{f(v)}dv-\frac{F(\hat{\theta})}{f(\hat{\theta})}\right]\frac{d\hat{\theta}(\alpha)}{d\alpha}-2(1-\alpha)\left[\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)^2+\int_{\hat{\theta}}^{\bar{\theta}}\frac{F(\hat{\theta})}{f(v)}dv\right]+(1-\alpha)^2\left[2\left(\frac{a}{(1-\alpha)}-\hat{\theta}\right)\frac{a}{(1-\alpha)^2}\right]$$

$$=(1-\alpha)^2\left[-\frac{F(\bar{\theta})}{f(\bar{\theta})}\right]\frac{d\hat{\theta}(\alpha)}{d\alpha}$$

Then $\lim_{\alpha\to\bar{\alpha}^-}SW'(\alpha)=-\frac{1}{(1-\alpha)^2}\left\{\left(b-\frac{\bar{\theta}}{a}\right)f(\bar{\theta})\int_{\underline{\theta}}^{\bar{\theta}}F(\theta)d\theta+\frac{1}{a}\left(\int_{\underline{\theta}}^{\bar{\theta}}F(\theta)d\theta\right)\right\}<0$.

These intermediate results together establish that the optimal degree of intervention lies on $(\underline{\alpha},\bar{\alpha})$, inducing a partially-separating menu as the market outcome. □

Maximum Information-Acquisition Policy

To investigate how the presence of an upper bound on information acquisition affects the vendor's optimal allocations, we again apply the Lagrange multiplier $\lambda_2(\theta)$ to the inequality presented in equation (4.5). The corresponding Hamiltonian-Lagrangian changes to $L(I,\eta,u,p_1,p_2,\lambda,\theta)\equiv H+\lambda_1(I(\theta)-\eta(\theta))+\lambda_2\frac{\theta}{\theta-a}u$.

The necessary conditions remain the same as in the benchmark case except that $p_1^*=p_1+\theta\lambda_1$ and $p_2^*=p_2+(a-\theta)\lambda_1-\lambda_2$. Accordingly, the counterpart of equation (14) is

$$\begin{cases}p_1(\theta)=\int_{\underline{\theta}}^{\theta}\{(b-(I(v)-\eta(v)))f(v)+\lambda_1(v)\}dv-\theta\lambda_1(\theta)\\ p_2(\theta)=\int_{\underline{\theta}}^{\theta}\{(I(v)-\eta(v)-b+1)f(v)-\lambda_1(v)\}dv+(\theta-a)\lambda_1(\theta)+\lambda_2(\theta)\end{cases}.$$

Analogously, we can derive some intermediate results that help characterize the optimal menu based on the necessary conditions. Particularly,

$$(\theta-a)p_1(\theta)+\theta p_2(\theta)=a\int_{\underline{\theta}}^{\theta}\left\{\left(I(v)-\eta(v)-\left(b-\frac{\theta}{a}\right)\right)f(v)-\lambda_1(v)\right\}dv+\theta\lambda_2(\theta) \tag{17}$$

Moreover, the transversality conditions imply $1+\lambda_2(\bar{\theta})=a\lambda_1(\bar{\theta})>0$, suggesting that the participation constraint should bind on optimality.

Proof of Lemma 4:

The proof proceeds in two steps. First, we show that full separation is no longer optimal when $K<\tilde{I}(\underline{\theta})$.

Suppose on the contrary, i.e. full separation is still optimal, we should have $(\theta-a)p_1+\theta p_2=0$ over $[\underline{\theta},\bar{\theta}]$. Governed by $u<0$, $I(\theta)\le K$ can only bind at $\underline{\theta}$, implying that $\lambda_2(\theta)$ is constant on $(\underline{\theta},\bar{\theta}]$. Using λ_2 to denote the constant value of $\lambda_2(\theta)$ on $(\underline{\theta},\bar{\theta}]$, we obtain:

$$\left(\theta-a\right)p_1(\theta)+\theta p_2(\theta)=a\int_{\underline{\theta}}^{\theta}\left\{I(v)-\eta(v)-\left(b-\frac{\theta}{a}\right)\right\}f(v)dv+\theta\lambda_2(\theta)$$

$$=a\int_{\underline{\theta}}^{\theta}\left\{I(v)-\eta(v)-\left(b-\frac{1}{a}\left(v+\frac{F(v)+\lambda_2}{f(v)}\right)\right)\right\}f(v)dv+\underline{\theta}\lambda_2=0 \quad \forall\theta\in(\underline{\theta},\overline{\theta}],$$

which implies $I^*(\theta)-\eta^*(\theta)=b-\frac{1}{a}\left(\theta+\frac{F(\theta)+\lambda_2}{f(\theta)}\right)$ and $\lambda_2=0$ over $[\underline{\theta},\overline{\theta}]$.

Hence, $u^*(\theta)=-\frac{\theta-a}{a^2}\left(\theta+\frac{F(\theta)}{f(\theta)}\right)'<0$, the same as in the benchmark case; so is $\left\{I^*(\theta),\eta^*(\theta)\right\}$ (because the participation constraint also binds at $\overline{\theta}$ as it does in the benchmark case). A contradiction is derived because $I^*(\theta)=\tilde{I}(\underline{\theta})>K$.

Second, we show that the principal's optimal menu will exhibit bunching at most on an interval at the bottom of the type space.

Suppose that $(\theta',\theta'')\subset[\underline{\theta},\overline{\theta}]$ is the first interval from below over which the optimal allocation is full separation (define θ' as $\theta'=\inf\left\{t:u^*(\theta)<0 \text{ for all } t<\theta<\theta''\right\}$ and θ'' as $\theta''=\sup\left\{t:u^*(\theta)<0 \text{ for all } \theta'<\theta<t\right\}$). Because u^* is discontinuous at θ', $\lambda_2(\theta)$ is continuous at this point, so is $\left(\theta-a\right)p_1(\theta)+\theta p_2(\theta)$ from equation (17). Hence we should have $\left(\theta'-a\right)p_1(\theta')+\theta' p_2(\theta')=0$. Moreover, given that $\left(\theta-a\right)p_1(\theta)+\theta p_2(\theta)=0$ on (θ',θ'') and $\lambda_2(\theta)$ is constant over $(\theta',\overline{\theta}]$, our aim is to show that θ'' should coincide with $\overline{\theta}$.

Since $\left(\theta-a\right)p_1(\theta)+\theta p_2(\theta)=0$ over (θ',θ''), $I(\theta)-\eta(\theta)=b-\frac{1}{a}\left(\theta+\frac{F(\theta)+\lambda_2}{f(\theta)}\right)$ $\forall\theta\in(\theta',\theta'')$.

Hence on (θ',θ''), $u^*(\theta)=-\frac{\theta-a}{a^2}\frac{\partial}{\partial\theta}\left(\theta+\frac{F(\theta)+\lambda_2}{f(\theta)}\right)$ and $\frac{\partial}{\partial\theta}\left(\theta+\frac{F(\theta)+\lambda_2}{f(\theta)}\right)>0$ (by the definition of (θ',θ'')). From Lemma A2, we know that $\frac{\partial}{\partial\theta}\left(\theta+\frac{F(\theta)+\lambda_2}{f(\theta)}\right)$ remains positive for all values of θ larger than θ''.

Therefore, if θ'' is different from $\overline{\theta}$, and $I(\theta)-\eta(\theta)-\left(b-\frac{1}{a}\left(\theta+\frac{F(\theta)+\lambda_2}{f(\theta)}\right)\right)$ starts to become positive from θ''^{+}, it remains so until $\overline{\theta}$ (because $u^*=0$ afterwards). We rewrite the expression of $p_1(\overline{\theta})$ as follows:

$$p_1(\overline{\theta})=\int_{\underline{\theta}}^{\overline{\theta}}\left(b-(I(v)-\eta(v))\right)f(v)dv-\overline{\theta}\lambda_1(\overline{\theta})$$

$$=-\frac{1}{a}\left(\left(\theta''-a\right)p_1(\theta'')+\theta''p_2(\theta'')\right)-\int_{\theta''}^{\overline{\theta}}\left\{I(\theta'')-\eta(\theta'')-\left(b-\frac{1}{a}\left(v+\frac{F(v)+\lambda_2}{f(v)}\right)\right)\right\}f(v)dv$$

$$=-\int_{\theta''}^{\overline{\theta}}\left\{I(\theta'')-\eta(\theta'')-\left(b-\frac{1}{a}\left(v+\frac{F(v)+\lambda_2}{f(v)}\right)\right)\right\}f(v)dv<0$$

The second equation follows from the fact that $(\theta''-a)\,p_1(\theta'')+\theta'' p_2(\theta'')=0$, and the last inequality follows from the fact that $\theta+\frac{F(\theta)+\lambda_2}{f(\theta)}>\theta''+\frac{F(\theta'')+\lambda_2}{f(\theta'')}$ over $(\theta'',\overline{\theta}]$. It contradicts the transversality conditions that $p_1(\overline{\theta})=0$. Hence, θ'' should coincide with $\overline{\theta}$, indicating that bunching can only occur for the lower interval of the type space, i.e. $[\underline{\theta},\theta']$.
Combining these two steps, we complete the proof for Lemma 4. □

Proof of Proposition 5:
From Lemma 4, the principal bunches the low-end of the type space $[\underline{\theta},\theta']$ (here again, θ' refers to the same concept as $\hat{\theta}$ in the main text does). Equation (4.6) is derived from $(\theta'-a)\,p_1(\theta')+\theta' p_2(\theta')=0$, and equation (4.7) results from the binding of the participation constraint at $\overline{\theta}$ and the binding of the information boundary constraint on $[\underline{\theta},\theta']$. □

Proof of Corollary 2:
1. According to the implicit function theorem, the following two equations are derived from equations (4.6) and (4.7):

$$-\left(\hat{\theta}+\frac{F(\hat{\theta})+\lambda}{f(\hat{\theta})}\right)' F(\hat{\theta})\frac{\partial\hat{\theta}}{\partial K}+\left(\hat{\theta}-\frac{F(\hat{\theta})}{f(\hat{\theta})}\right)\frac{\partial\lambda}{\partial K}=0\,,$$

and $$-\hat{\theta}\left(\hat{\theta}+\frac{F(\hat{\theta})+\lambda}{f(\hat{\theta})}\right)'\frac{\partial\hat{\theta}}{\partial K}-\left(\frac{\hat{\theta}}{f(\hat{\theta})}+\int_{\hat{\theta}(K)}^{\overline{\theta}}\frac{1}{f(\theta)}d\theta\right)\frac{\partial\lambda}{\partial K}=1.$$

From them, $$\frac{\partial\hat{\theta}}{\partial K}=\frac{-\left(\hat{\theta}-\frac{F(\hat{\theta})}{f(\hat{\theta})}\right)}{\left(\hat{\theta}+\frac{F(\hat{\theta})+\lambda}{f(\hat{\theta})}\right)'\left(\hat{\theta}^2+F(\hat{\theta})\int_{\hat{\theta}(K)}^{\overline{\theta}}\frac{1}{f(\theta)}d\theta\right)}.$$

It can be verified that both the denominator and the nominator are positive ($\lambda=\frac{F(\hat{\theta})^2}{\hat{\theta}f(\hat{\theta})-F(\hat{\theta})}>0$ from equation (4.6) and the definition of λ).

2. From equation (4.6) and the definition of λ, $\hat{\theta}f(\hat{\theta})-F(\hat{\theta})>0$.

If $\theta f(\theta)-F(\theta)>0$ on the entire type space, $\varnothing=\inf\left\{\theta:\theta-\frac{F(\theta)}{f(\theta)}=0\right\}$, and the upper bound of $\hat{\theta}(K)$ is $\overline{\theta}$. We shall demonstrate that if $\theta f(\theta)-F(\theta)\le 0$ at some $\tilde{\theta}\in[\underline{\theta},\overline{\theta}]$, $\theta f(\theta)-F(\theta)\le 0$ on the remaining type space, i.e. $(\tilde{\theta},\overline{\theta}]$.

First, the first-order derivative of $\theta f(\theta)-F(\theta)$ is $\frac{d}{d\theta}\left[\theta f(\theta)-F(\theta)\right]=\theta f'(\theta)$. It implies that the monotonicity of $\theta f(\theta)-F(\theta)$ is the same as that of $f(\theta)$: it first starts with a non-decreasing interval, and then is reversed to be non-increasing after some tipping point (this is due to the unimodality of $f(\theta)$). Since $\underline{\theta}f(\underline{\theta})-F(\underline{\theta})>0$, $\theta f(\theta)-F(\theta)$ equals to 0 only on its non-increasing interval. The result follows directly from the uni-modality of the density function.

Moreover, equation (4.6) implies that $\frac{d\lambda}{d\hat{\theta}} > 0$ on $[\underline{\theta}, \tilde{\theta})$, and increases to infinity when $\hat{\theta}$ approaches to $\tilde{\theta}$. Therefore, for a log-concave distribution with $\inf\{\theta : \theta f(\theta) - F(\theta) = 0\} < \overline{\theta}$ satisfied, the bunching region never extends beyond $\tilde{\theta} \equiv \inf\left\{\theta : \theta - \frac{F(\theta)}{f(\theta)} = 0\right\}$ for any arbitrarily low boundary (this assertion also requires full participation on optimality, i.e., $I^*(\theta) > 0$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$). As a consequence, a θ belonging to the top interval of the support cannot be an eligible candidate for $\hat{\theta}$.

Proof of Proposition 6:

For the set of partially-separating menus, we can show that

$$SW'(K) = \frac{1}{a} \frac{1}{\left(\hat{\theta}^2 + F(\hat{\theta}) \int_{\hat{\theta}(K)}^{\overline{\theta}} \frac{1}{f(\theta)} d\theta\right)} \left[\hat{\theta} \int_{\underline{\theta}}^{\hat{\theta}(K)} \left(\frac{1}{a}\left(\hat{\theta} + \frac{F(\hat{\theta}) + \lambda}{f(\hat{\theta})}\right) - \frac{\theta}{a}\right) f(\theta) d\theta + F(\hat{\theta}) \int_{\hat{\theta}(K)}^{\overline{\theta}} \left(\frac{1}{a}\left(\theta + \frac{F(\theta) + \lambda}{f(\theta)}\right) - \frac{\theta}{a}\right) f(\theta) d\theta\right] > 0$$

for any $K < \tilde{I}(\underline{\theta})$.

For the set of pure-bunching menus, again $SW'(K) > 0$. □