Application Letter for
"NBER Young Scholars Workshop on the Economics of Artificial Intelligence"

Dear Sir/Madam,

I am a PhD student from Rotman School of Management, University of Toronto. I'm writing to express my interest in participating in the NBER Young Scholars Workshop on the Economics of Artificial Intelligence. Since I live in Toronto, where the workshop will take place, I'm NOT applying for funding for my travelling and accommodation.

My interest in AI began with my concerns about the welfare implications of privacy regulations in recent years. Firms collect consumer data via their browsing, searching and even chatting history online to infer consumers' preference for products. The advancement of algorithm has allowed firms to send relevant product information precisely to their target consumers, and charge discriminating prices accordingly. At the same time, concerns for the abuse of private data have pushed many governments to implement privacy regulations. The most notable one is EU's "General Data Protection Regulation" that became in effect in 2018. However, such regulatory attempts to protect the privacy of consumers' data might backfire. One possible impact is that it may result in firms manipulating the quality of products they provide, including also ads and spams, so as to encourage more consumers to share their data with firms, if that's to the firms' interest. A more obvious impact is that such regulation would impede the improvement of algorithms that aims at providing consumers with better customized user experience. Impacts as what are in the examples above make it ambiguous whether these privacy regulations would actually enhance the social welfare.

Inspired by this concern, I wrote a research proposal regarding this issue as a course project, which I'd like to discuss with other participants of this workshop using this great opportunity. This proposal is attached following my resume.

I am also very interested in understanding the impact of the automation that replaces routine labour and in the future may replace even creative labour with the advancement of AI.

As such, I'd appreciate it a lot you could consider my application for participating in this workshop favourably.

Sincerely,

Ruoxi Xie

RUOXI XIE

PhD House, 105 St George, Ruoxi.Xie17@Rotman.Utoronto.Ca
M5S 3E6, Toronto, Canada +1 6478659135

DOCTORAL STUDIES

UNIVERSITY OF TORONTO Toronto, Canada, *2017-Present*
PhD in Economic Analysis and Policy

PRIOR EXPERIENCE

CENTRO DE ESTUDIO MONETARIOS Y FINANCIEROS(CEMFI) Madrid, Spain, *2017*
Master in Economics and Finance

XIAMEN UNIVERSITY Xiamen, China, *2015*
Bachelor of Arts in Economics

Exchange Studies

NATIONAL UNIVERSITY OF SINGAPORE Singapore, *2014*

RESEARCH INTERESTS

Industrial Organization, International Trade, Applied Economic Theory, Behaviroal Economics

KEY COURSEWORK

Advanced courses in Microeconomics, Macroeconomics and Econometrics; Advanced topics in Industrial Organization; Innovation and Economics; Empirical Applications of Economic Theory; Empirical Methods of Industrial Organization, etc.

TEACHING EXPERIENCE

Teaching assistant for MBA Courses "Decision Making with Models & Data", and "Managerial Economics". *2018*
Teaching assistant for undergraduate course "Game Theory for Business Strategy". *2019*

RESEARCH EXPERIENCES

Research assistant for Professor Bernardo Blum, University of Toronto. *2018-2019*
Research assistant for Professor Guillermo Caruana, CEMFI. *2016*
Research assistant in Financial and Economics Experimental Laboratory, Xiamen University. *2014*

LANGUAGE SKILLS Chinese (Native) , English (Fluent) , Spanish (Beginner)

COMPUTER SKILLS STATA, Matlab, LaTex, Microsoft office

GRANTS & SCHOLARSHIPS & AWARDS

Robert House Fellowship *2018-Present*
University of Toronto PhD Scholarship *2017-Present*
University of Toronto Director's Fellowship *2017-Present*
CEMFI Master Full Scholarship *2015-2017*
Best Undergraduate Thesis Award *2015*
Outstanding Graduates Award *2015*
TF-LEaRN Scholarship from Temasek Foundation *2014*
Outstanding Academic Performance Scholarship *2013*
National Scholarship *2012*

Privacy Regulation and Advertisement Design with Heterogeneous Valuation for Match

Ruoxi Xie

April 2019

1 Introduction

It is widely acknowledged that firms now collect consumer data via their browsing, searching and even chatting history online to infer consumers' preference for products. This enables firms to send relevant product information precisely to their target consumers, and charge discriminating prices accordingly. However, this triggers concerns for privacy disclosure. People intrinsically value privacy. Also, some might worry about being price discriminated against. From a regulatory point of view, the impact of this data collection and utilization can have two effects. Firstly, it improves efficiency in the sense that for a given space of information slot, it can be targeted toward those who value it the most. On the other hand, everyone suffers a loss from having their privacy disclosed, especially for those that do not enjoy much from the efficiency gain.

In recent years, concerns for the abuse of private data have pushed many governments to implement privacy regulations. The most notable one is EU's "General Data Protection Regulation" that became in effect in 2018. This law aims at giving control to individuals over their personal data.

The goal of this study is to model how regulatory attempts to protect the privacy of consumersâ data affect firms' incentive to provide clear information in their advertisement and the welfare implication of this impact. In particular, it attempts to show how it is possible that privacy regulation that aims at promoting social welfare could in effect hurt consumers under some conditions, even when consumers optimally choose to share or not to share their personal information with firms.

The modelling part of this proposal is very parsimonious and ad-hoc. It only aims at delivering the rough idea that it is possible that privacy regulation could trigger backlash in welfare loss. It considers a special case in which firms can utilize the clarity level of their recommendation to manipulate consumers' incentive to share their information. In particular, firms optimally manipulate product recommendations to persuade consumers whose personal information is of high value to them to share their data with them. It characterizes some aspects of the equilibrium and involves some discussion of the intuition

of the equilibrium but does not involve rigorous proofs. More rigorous and generalizable modelling is left for the future.

2 Literature Review

This study is related to several strands of literature.

In terms of a focus on privacy and efficiency, the literature dates back to Spence (1978) and Posner (1981). Posner (1981) argues that allowing individuals to conceal their personal information is detrimental to efficiency because it leads to inefficient decisions made by the uninformed agents. Spence (1978) studies a model where agents can signal their types. It shows that there are situations where cost from allowing signalling outweigh benefit. In this proposed work, consumers choice to share or not share their personal information can be viewed as a signalling activity to some degree. It delivers payoff-relevant information about individuals and firms' choices of price and product recommendation are affected by this signalling activity. More recently, Calzolari and Pavan (2006) studies the share of customer information between companies. It shows that in some cases this share of information enhances social welfare. Campbell et al. (2015) studies the impact of privacy regulation on market structure. It shows that when the opt-in consent as required by the regulation is costly for consumers, regulation will result in a more concentrated market because consumers prefer to opt-in to firms with a large scope of services.

This study allows firms to price discriminate using personal information they collect from consumers and consumers may strategically choose to disclose their information or not. This scenario is shared by many recent studies. Taylor (2004) studies the welfare comparison under the two settings, in one of them personal information is confidential and in another it is disclosed and firms can sell their information to other firms for price discrimination. It shows that the welfare comparison depends critically on consumers' level of sophistication. Naive consumers who are not aware of data tracking and price discrimination lose their surplus to firms under the disclosure regime, while sophisticated consumers would adjust their behaviors accordingly to avoid such surplus extraction. For this reason, firms would voluntarily protect consumers' privacy and thus regulation is redundant. Villas-Boas (2004) shows that the ability to implement dynamic targeted pricing can actually lead to lower profit to a monopolist when consumers are strategic inter-temporarily. **?** studies a two-period model where merchants have tracking technology and consumers can anonymize themselves. They find that the merchant would find it profitable to track consumers only when tracking can also enhance personalized services. Ichihashi (2019) studies the case where consumers trade-off between a more accurate product recommendation from a multi-product seller and price discrimination. It demonstrates that the seller prefers to commit to not use information for pricing to encourage information disclosure, but this commitment hurts consumers.

This study differs from the literature in that firms not only decide on what prices to set and and what products to recommend, but also on the clarity level of their recommendation. Less clarity can lead to consumers' misunderstanding of the match between

them and the product, such that sometimes consumers can end up buying products that are of no value to them. In particular, in this study, lower clarity level is used by sellers to discourage consumers from opt-out from disclosing their personal information.

3 Model

3.1 Market for Products

Consider an online market with a monopoly and a continuum of consumers with unit demand. The assumption of monopoly is a simplification for highlighting the main points. It can be think of as an integrated entity composed of numerous sellers and a platform, whose joint goal is to maximize the total profit they receive from selling products to the consumers and then the sellers and the platform will share the maximized profit ex-post according to some rule they agree to ex-ante. Each product j has a horizontal characteristics. Consumers are heterogeneous along two dimensions: the match between their taste and the product's horizontal characteristics: m_{ij}, and their valuation for the match between their taste and the products' horizontal characteristics: θ_i. The monopoly has a sufficiently large variety of products that it can provide to each consumer i products of any $m_{ij} \in (0, \overline{m}]$. θ_i is i.i.d in [0, 1] according to cdf $F(\theta)$. Besides choosing m_{ij}, the firm can also choose an information clarity level $\alpha_{ij} \in [0, 1]$ when it recommends a product to consumers. For product with type (α_{ij}, m_{ij}), it means that product j is of match value m_{ij} to the consumer i with probability α and of match value 0 with probability $1-\alpha$.Product cost is normalized to 0.

3.2 Privacy

θ_i is ONLY known to consumer i. If a consumer's private information is disclosed, m_{ij} is known by the monopoly. This scenario corresponds to the case where the platform can formulate consumers' favorite product according consumers' browsing and searching history but is not able to know precisely a consumer's willingness to pay for a product. If a consumer's private information is not disclosed, m_{ij} is not known by the monopoly. Consumers suffer from a privacy disclosure loss of s.
There is a regulator that governs the market. If the regulator imposes privacy regulation, the monopoly is required to give consumers the option to opt-out from disclosing their information. Otherwise, all consumers' information is disclosed by default.

3.3 Timing

1. Regulator decides to impose privacy regulation or not.
2. If regulator decides to regulate, consumers decide whether or not to opt-out.
3. Monopoly chooses price, information clarity level and horizontal characteristics for each consumer.
4. Consumers observe price, information clarity level and horizontal characteristics and make purchase decisions.

3.4 Demand

Each consumer is only aware of the product that the monopoly recommends/shows to them. When product of type (α_{ij}, m_{ij}) is recommended to consumer i and price is p_{ij}, the consumer buys the product iff $\theta_i \alpha_{ij} m_{ij} \geq p_{ij} \iff \theta_i \geq \frac{p_{ij}}{\alpha_{ij} m_{ij}} \equiv \tilde{\theta}$. Consumer i's utility is

$$U_{ij} = \begin{cases} -s, & \text{if information is disclosed and not buy,} \\ \theta_i m_{ij} - p_{ij} - s, & \text{if information is disclosed, buy \& match value is } m_{ij}, \\ -p_{ij} - s, & \text{if information is disclosed, buy \& match value is 0,} \\ 0, & \text{if information is not disclosed and not buy,} \\ \theta_i m_{ij} - p_{ij}, & \text{if information is not disclosed, buy \&match value is } m_{ij}, \\ -p_{ij}, & \text{if information is not disclosed, buy \& match value is 0,} \end{cases} \tag{1}$$

3.5 Supply Without Privacy Regulation

$$\max_{p_{ij}, \alpha_{ij}, m_{ij}} p_{ij}(1 - F(\tilde{\theta})) \tag{2}$$

Trivially, $m_{ij}^* = \overline{m}$ and $\alpha_{ij} = 1$ and p_{ij}^* is characterized by

$$F(\frac{p_{ij}^*}{\overline{m}}) + \frac{p_{ij}^*}{\overline{m}} f(\frac{p_{ij}^*}{\overline{m}}) = 1 \tag{3}$$

Note that in this case, every consumer consumes the product with their favorite horizontal characteristics and there is full clarity in the product recommendation. In addition, they are all charged the same price $p_{ij}^* = p^*$ that is characterized by equation (3).

3.6 Total Welfare Without Privacy Regulation

$$\begin{aligned} W^{NR} &= \int_{\frac{p^*}{\overline{m}}}^{1} (\theta \overline{m}) dF(\theta) - s \\ &= \overline{m} E(\theta | \theta > \frac{p^*}{\overline{m}})(1 - F(\frac{p^*}{\overline{m}})) - s \end{aligned} \tag{4}$$

3.7 Demand With Privacy Regulation

Now, consumers are given the option to opt out from sharing their information. For consumers that opt-out, the monopoly cannot customize their product. Therefore they randomly choose horizontal characteristics to provide to each consumer and thus monopoly (consumer) expects to provide(receive) a match value of $E(\theta)$. Since the profit maximization is the same for all consumers that opt out, they will all be charged with the same price p^{out} and provided with the same level of clarity α^{out} in the recommendation. Similarly, all consumers that choose to share their information with the monopoly will be charged with the same price p^{in}, provided with the same level of clarity α^{in} and same match level m^{in}. The consumer type $\hat{\theta}$ that is indifferent between opt-in and opt-out is

determined by:

$$\begin{aligned} \hat{\theta}\alpha^{in}m^{in} - p^{in} - s &= \max\{0, \hat{\theta}\alpha^{out}E(m) - p^{out}\} \\ \iff \hat{\theta} &= \max\{\frac{p^{in}+s}{\alpha^{in}m^{in}}, \frac{p^{in}-p^{out}+s}{\alpha^{in}m^{in}-\alpha^{out}E(m)}\} \end{aligned} \tag{5}$$

Opt-in Decision: Opt-in iff $\theta_i \geq \hat{\theta}$.
Purchase decision:
Conditional on opt-in, consumers with $\theta_i \geq \frac{p^{in}}{\alpha^{in}m^{in}} \equiv \tilde{\theta}^{in}$ will buy. Conditional on opt-out, consumers with $\theta_i \geq \frac{p^{out}}{\alpha^{out}E(m)} \equiv \tilde{\theta}^{out}$ will buy.
Note that $\hat{\theta} > \tilde{\theta}^{in}$, which means that every consumer that opt-in will buy. Because consumers that opt-in can rationally anticipate to receive α^{in} and m^{in} that are chosen by the monopoly in equilibrium and there is privacy disclosure loss when opt-in, consumers would only opt-in when they expect to buy.

4 Supply With Privacy Regulation

Monopoly's maximization problem:

$$\max_{p^{in},p^{out},\alpha^{in},\alpha^{out},m^{in}} p^{in}Pr(\theta \geq \hat{\theta}) + p^{out}Pr(\theta \geq \tilde{\theta}^{out}|\theta < \hat{\theta})Pr(\theta < \hat{\theta}) \tag{6}$$

Again, it is easy to see that $m^{in^*} = \overline{m}$ and $\alpha^{in^*} = 1$. Suppose not, increasing m^{in} and α^{in} a little and increasing α^{out} a little correspondingly, keeping every else constant, could keep $\hat{\theta}$ unchanged. And since m^{in}, α^{in} and α^{out} increases, demand increases for both opt-in and opt-out groups. Therefore profit will increase. Hence it has to be that $m^{in^*} = \overline{m}$ and $\alpha^{in^*} = 1$.

Moreover, $p^{in*} \geq p^*$. This is because those who opt in is a selective group of all consumers in the sense that they have a higher θ^i than those who opt-out, thus a lower price elasticity. An immediate result from $p^{in*} \geq p^*$is that $\hat{\theta} \geq \tilde{\theta}$. What's more, even if $p^{in} = p^*$, those who obtain an expected payoff from buying the product less than s would opt-out in the regulated case. Therefore $\hat{\theta} > \tilde{\theta}$. In another word, fewer consumers are enjoying customized horizontal characteristics in the regulated environment than in the non-regulated case.

Notably, $\alpha^{out^*} \in (0, 1)$.
$\alpha^{out^*} > 0$ because in equilibrium, consumers with low θ will opt-out. Suppose $\alpha^{out^*} = 0$, it is always profitable to increase α^{out} a little and charge a price low enough that some of these consumers will buy but consumers that opt-in wouldn't want to deviate.
$\alpha^{out^*} < 1$ for similar reason as the "distortion at the bottom" result of the classic menu-pricing. Consumers with a very large θ wouldn't want to opt-out as long as p^{in} is not too big. Consumers with a very low θ wouldn't opt-in as long as p^{out} is not extremely low. It is those consumers that have an intermediate level of θ that the monopoly want to convince them to opt-in and extract a higher rent from them through customization. However they need to compensate for their privacy loss and higher price. Consider a

marginal consumer. The monopoly could either lower p^{in} or increase p^{out} or lower α^{out} to achieve the goal. They wouldn't want to lower p^{in} because all the price charged to those opt-in would be lower for this marginal consumer. The monopoly would also prefer decreasing α to increasing p^{out} because this way it loses fewer consumers with low θ.

5 Welfare Analysis

The comparison between the welfare under regulation and non-regulation requires a comparison between the efficiency loss and privacy protection gain.
The efficiency loss are two-folds. Firstly, fewer consumers are enjoying customized product. Second, The monopoly optimally choose a clarity level lower than 1 for consumer that opt-out to induce consumers with a relatively higher valuation for match to opt-in. This lower clarity level leads to efficiency loss because some consumers end up being matched with goods that give them a value of 0.
Those who opt-out now enjoy a privacy protection. Total privacy protection gain is $\hat{\theta}s$.
A precise welfare analysis require solving the equilibrium full and perhaps a specification of distributions. But intuitively, while privacy protection gain is linear in $\hat{\theta}$, efficiency loss is not due to the none linear relationship between $\hat{\theta}$ and α as shown in equation (5). Therefore, it is possible that under some conditions, the welfare loss under privacy regulation can exceed the gain from privacy protection.

References

Calzolari, G. and Pavan, A. (2006). On the optimality of privacy in sequential contracting. *Journal of Economic theory*, 130(1):168–204.

Campbell, J., Goldfarb, A., and Tucker, C. (2015). Privacy regulation and market structure. *Journal of Economics & Management Strategy*, 24(1):47–73.

Ichihashi, S. (2019). Online privacy and information disclosure by consumers. *Available at SSRN 3112905*.

Posner, R. A. (1981). The economics of privacy. *The American economic review*, 71(2):405–409.

Spence, M. (1978). Job market signaling. In *Uncertainty in Economics*, pages 281–306. Elsevier.

Taylor, C. R. (2004). Consumer privacy and the market for customer information. *RAND Journal of Economics*, pages 631–650.

Villas-Boas, J. M. (2004). Price cycles in markets with customer recognition. *RAND Journal of Economics*, pages 486–501.