Privacy on a platform[*]

Danny Kwon[†]

Abstract

We review the current literature on privacy and how it models agent behaviour on platforms. We develop a model which highlights agent behaviour when faced with privacy decisions on a monopoly digital platform. We find the profit maximising privacy setting for developers, along with the profit maximising privacy standard that the platform can enforce, and characterise the appropriate conditions required for the privacy standard to be binding. We also perform comparative statics to identify how the size of relevant variables affects welfare outcomes. Additionally, we determine the welfare maximising level of privacy, and simulate scenarios to identify when the welfare maximising privacy setting can be above, below, or equal to the profit maximising privacy values.

Keywords: *privacy, platform, app development, quality*

JEL classification: *D4, L2, L86*

[*] Declarations of interest: None.

[†] Corresponding author. Department of Economics and Centre for Mathematical Social Science, The University of Auckland, Sir Owen G Glenn Building, 12 Grafton Rd, 1010 Auckland, New Zealand. E-mail: d.kwon@auckland.ac.nz.

1 Introduction

1.1 Context

In January 2021, WhatsApp, the world's largest digital messaging application (WhatsApp, 2020), revised their user privacy policy (WhatsApp, 2021). The change was aimed at the data management methods for businesses who use WhatsApp as their primary online communication tool, giving them the choice to store their user activity logs on Facebook. For a typical user, nothing was fundamentally affected, and yet the rumour of their own private data potentially being signed over to Facebook was enough to trigger massive outrage, with users moving to alternative messaging applications, such as Signal and Telegram, both aimed at providing a more private experience (Parikh, 2021).

This was not the first time a digital platform has been the centre of a privacy related controversy, nor will it be the last. The UK COVID trial data breach in 2021, private Zoom meeting breaches in 2020, the Facebook–Cambridge Analytica data scandal revealed in 2018, the 2017 Equifax hack, the Experian databreach in 2015, the 2014 iCloud breach (dubbed the Fappening), the 3 billion Yahoo account breach in 2013; privacy concerns have plagued users of digital platforms since internet access became widespread, and the number of events just continues to increase.

Furthermore, these digital privacy issues have had impacts that far exceed the users themselves. The Cambridge Analytica incident affected the U.S. 2016 election, culminating with Trump winning along with potentially affecting the outcome of the Brexit referendum. These effects have forced policymakers into action, with mixed results. Trump's restriction of American technologies being used by Chinese firms, with the prime example being HuaWei restricted from using the most ubiquitous Android OS for their smartphones, has had rippling effects, with privacy concerns suddenly becoming an issue of national security.

Furthermore, this fight for privacy has also become a marketable point for digital firms. Apple has begun marketing itself as a "private" firm, with quotes such as "Privacy is a fundamental human right" (Apple, 2021), highlighting how their smartphones allow a higher degree of privacy, through strict monitoring of applications on their own App Store, while also simultaneously providing a high level of security for personal data.

Other firms argue that users should be willing to give data to them for optimisation of services. Google does this often, promoting that giving them your own personal data allows for optimisation of their services for you, generating greater value (Rosenberg, 2019). This has been called into question, as the data captured appears to far exceed any business need, but is rather captured for capture's sake.

Finally, there are digital firms that place little emphasis on privacy, preferring their stance to be as unclear as possible. Facebook is the prime example, with privacy guidelines so ambiguous that interpretation is close to impossible. Rather than selling the idea that they capture user data, they instead push the focus onto user-to-user interactions, highlighting how it is possible to limit interaction between them, thus avoiding the topic entirely.

A notable pattern emerges, as at the very centre of privacy exist digital platforms. There is a fundamental link between how a digital platform operates and the level of privacy it is willing to give to users. Thus, it is worthwhile considering how these privacy settings will affect not just users, but potential producers in the form of application developers as well.

1.2 Research question

How do platform specific privacy settings affect agent equilibrium participation, pricing, and privacy levels in a digital platform market?

1.3 Literature review

The research on privacy and data control on digital platforms is a growing space. While privacy has been discussed in economics since the early 1960s with Stigler (1961), most of the research was approaching it from an asymmetric information perspective. After this very early work, a "First Wave" of privacy literature was produced between the 1970s and 1980s (Acquisti et al., 2016). This mostly encompassed qualitative studies, highlighting the need for further research into privacy (Posner, 1981) given the rapid regulations that were being developed, along with how an individual who has higher degrees of privacy can have negative economic impacts through inefficient allocation of labour resources (Stigler, 1980). However, this devaluation of privacy was challenged by Hirshleifer (1980) who points out that this oversimplification is a harmful rhetoric with future ethical implications.

After the first wave, privacy literature enters another slump. That is, until the proliferation of digital technologies in the 1990s. With household computing and internet access becoming more commonplace, privacy comes forth once more as a concern. As before, most of the research focuses on more qualitative arguments, rather than foundational quantitative theory. Varian (2002) highlights some of the earlier issues with privacy in the "information age", highlighting how one's attention is valuable, and proposing that privacy rights are equivalent to the "right not to be annoyed". He develops this further, proposing potential ways that privacy could be valued in the future along with future policy issues that may come about. Noam (1997) similarly predicted current privacy issues, pointing out that consumers do indeed care, and want privacy, and will self select into privacy levels that they prefer. Furthermore, he identifies that regulatory regimes will be the deciding factor in whether consumers are paid to give up their privacy, or pay to protect themselves.

Jumping to the present day, the research space around privacy continues to expand. Rather than arguments regarding the relevance that privacy has, theoretical models are proposed which attempt to model how society as a whole interacts with this idea of privacy. Choi et al. (2019) provides a model which highlights how consumer data is over-extracted when facing a monopoly content provider (implying privacy rights being abused). This reflection of the modern day digital ecosystem is quite apt, with many websites being the sole provider of such content (such as Facebook for social media). Board and Lu (2018) also demonstrates how information exchange can work in inverse, with buyers seeking information regarding sellers, turning the standard privacy paradigm upside down.

Research also analyses how policy affects consumer privacy. Campbell et al. (2015) highlight how implementation of privacy policy can potentially be anti-competitive, and that this effect is magnified online, given the large number of entrenched incumbent firms.

Acemoglu et al. (2019) integrates the idea of privacy with digital platforms highlighting how platforms can use data to identify the typing of its users, finding that the presence of platforms lead to data sharing inefficiencies from the user side, with them oversharing information and data. There is also a reasonable amount of research with respect to platforms acting as data intermediaries. Hagiu and Jullien (2011) demonstrate why these intermediaries would exist in the first place, with Bergemann and Bonatti (2015) supporting this conclusion through their model in which a data provider intermediary leads to more efficient outcomes, but also has the ability to control data flows to ensure a more profitable environment. This continues with Yang (2020) demonstrating the interactions for a single data broker, and how that leads to zero consumer surplus through intense price discrimination.

This continues with works theorising how digital businesses characterise privacy. Fainmesser et al. (2019) provides a framework which demonstrates how certain qualities of businesses can lead to different data storage methodologies, and thus provide differing levels of privacy. Bergemann et al. (2020) studies the other side, demonstrating why and how consumers would voluntarily provide information to firms and how certain policy settings regarding property rights over data can affect allocative efficiency. Ichihashi (2020) provides a novel addition, demonstrating how users interact with the platform itself given platform privacy settings, finding that activity increases as users expect higher levels of privacy, or as users lose all of their privacy. It also highlights how platforms do indeed profit from the user data they capture through user activity.

Fundamentally, the literature highlights two major aspects of privacy. First, definitions are not all aligned, with authors disagreeing with respect to what privacy exactly is, and second, individuals greatly value privacy. While the individual user may have a different definition of what privacy is compared to another, it is agreed that privacy is valuable, and by extension, a user's own personal data is also valued.

1.4 The definition of privacy

It should be noted that throughout the literature, there is yet to be one cohesive definition of privacy. Whether privacy is defined as "protection against access" or "control over usage" (Acquisti et al., 2016) is contested, and while we approach privacy from a "control over usage" perspective, almost regarding it as an opaqueness value, i.e. as privacy levels increase, the ways in which a user's data will be used becomes better defined (thus giving the user more information to make a decision on whether to distribute that data), it need not only follow this definition. For a matter of simplicity, we will regard privacy as a numerical value that represents the opaqueness of usage, with higher values representing better defined usage, making it a vertically differentiated quality.

2 Basic Ingredients

Players We consider a model for a monopolist digital platform which connects platform application developers with platform users. There are three groups of agents:

- Platform: Let there be a monopolist which hosts and connects users with developers and profits from every user of an application through acquisition of data. The platform also decides the minimum standard of privacy ϕ_{min} that is followed by the hosted developers. This can be seen as an app store which is a platform that digitally distributes software to willing users, such as the Apple App Store, or the Google Play Store.

- Developers: There exist developers that produce applications to be hosted on the platform and consumed by users. These developers are characterised by their idiosyncratic development cost.

  Assumption 1. *There exists a continuum of developers a, of which the mass is normalised to one, where their development cost $\theta_a \in [\underline{\theta}, \bar{\theta}]$ is drawn from the uniform distribution function $g(\theta)$.*

  The developer chooses their pricing p_a, along with their privacy level ϕ_a. Developers profit from users consuming their application for a given price, and will pay costs of development, and privacy implementation if choosing to participate and develop.

- Users: There exists users potentially interested in gaining access to the developed applications present on the platform. Each user possesses an idiosyncratic preference over the desired minimum privacy level for applications, denoted by ϕ_i.

  Assumption 2. *There exists users i, of which the mass is normalised to one, where their minimum desired level of privacy demanded $\phi_i \in [\underline{\phi}, \bar{\phi}]$ is drawn from the uniform distribution function $f(\phi)$.*

  Users can consume as many or as few applications as they desire, and make the decision to consume applications based on the intrinsic utility of the application, the price, the privacy level of the application, and their own privacy preference.

A subtle outcome of these agent settings is that a single user i can be counted as more than a "single user". This is due to the fact that the platform sees each user for each application as unique: user i using applications a and b would be counted as two users, with respect to each application.

We consider the best response functions of pricing and privacy for the developer, while also considering the best response minimum privacy setting that the platform will implement. We find equilibrium levels of user and developer participation, along with developer and platform profit.

The game The agents all engage in the privacy setting game, where they seek to maximise their objective function. While we analyse each of those objectives in detail in section 2.1, we highlight them below.

User i's utility function for application a is:

$$u_{i,a} = \omega - p_a + \phi_a - \phi_i.$$

Note that this function is specific for application a. Users will face a utility function for every single application that is actively offered and that users eventually choose to install for them to access/use.

Developer a's profit function is:

$$\pi_a = p_a \int_{\underline{\phi}}^{\widetilde{\phi}} f(\phi)\, d\phi - \theta_a - \frac{\phi_a^2}{2}.$$

The upper bound of the integral $\widetilde{\phi}$ denotes the marginal user for developer a. This will be further defined in section 2.1.

The platform's profit function is:

$$\Pi = \int_{\underline{\phi}}^{\widetilde{\phi}} f(\phi)\, d\phi \int_{\underline{\theta}}^{\widetilde{\theta}} g(\theta)\, d\theta.$$

The $\widetilde{\theta}$ upper limit for the integral of function $g(\theta_a)$ denotes the marginal developer who will still be producing an application.

Timing The privacy setting game proceeds as follows:

1. A minimum privacy standard is set by the platform. This standard must be met by participating developers.
2. Developers choose whether to participate and develop an application.
3. Developers set their privacy level.
4. Developers set their prices.
5. Users evaluate whether to participate.
6. Users then decide which apps to use on the platform.

2.1 Analysis

We begin by solving the game backwards. Proofs are relegated to a separate appendix.

Users A user i's utility for installing and using developer a's app is specified as:

$$u_{i,a} = \omega - p_a + \phi_a - \phi_i. \tag{1}$$

Note that users do not have a constraint placed on the number of apps they can install. Therefore, users choose to install any app that is associated with a non-negative net utility. This emulates the current ecosystem in which users are able to install multiple apps on their phone from their platform of choice, where each has a non-negative price (given different apps come at different prices). This assumes that applications are non-rival to each other: Time allocated to using an application does not reduce the time allocated to others. In reality, a user's time will be constrained, and the use of alternative apps is indeed competing, but we abstract from this and consider the option to use the app as the source of utility for users, i.e., regardless of whether the app gets used and how often. Therefore, we interchangeably refer to the term 'install' and 'use' throughout this analysis. While a user will not always be "using" an application, simply having it available as an option is enough to generate utility. As an example, while a user of Facebook may not always be "using" the application, having it downloaded and available for use on their smartphone is fundamentally generating utility as they have access to it at will.

The marginal user for an app is the user whose desired privacy level is just sufficient to make them indifferent between downloading and not downloading that app. More formally, it means that a user's individual rationality constraint (IR) is satisfied with equality, given the outcome of the alternative option, not installing that app, is equal to zero. Thus, this defines $\widetilde{\phi}$, the level of desired privacy that separates the active and inactive users, as a function of each price and privacy level chosen by each app a.

Developer a's profit function is specified as follows, using the threshold of desired privacy preferences discriminating between active and non-active users of that app, as defined by the (IR) discussed above:

$$\pi_a = p_a \int_{\underline{\phi}}^{\omega - p_a + \phi_a} f(\phi)\, d\phi - \theta_a - \frac{\phi_a^2}{2}. \tag{2}$$

where $\int_{\underline{\phi}}^{\omega - p_a + \phi_a} f(\phi)$ captures the mass of active users, i.e., choosing to install app a.

We take the partial derivative of the developer's profit function with respect to their price to obtain their best response price function.

Lemma 1. *The developer's best response pricing p_a is $\frac{\omega + \phi_a - \underline{\phi}}{2}$.*

Substituting best response pricing into the developer profit function, we take the partial derivative with respect to ϕ_a to identify the optimal privacy setting.

Lemma 2. *The developer's optimal privacy setting ϕ_a given its best response pricing will be $\omega - \underline{\phi}$.*

Comment 1. *Due to no actions being taken by other agents, the timing of setting prices and privacy levels for developers can be interchanged with no impact on results.*

From these results, we note that there is a natural restriction on the sizes of ω and $\underline{\phi}$. If $\omega < \underline{\phi}$, prices can turn negative, violating the non-negativity condition. This leads to the following assumption, to ensure that the market is viable:

Assumption 3. $\omega \geq \underline{\phi}$.

Note that the fringe case of $\omega = \underline{\phi}$ will cause developers to never participate on the platform, thus, while assumption 3 is necessary, it is not sufficient. For the rest of the analysis, we will implicitly assume that $\omega > \underline{\phi}$.

Assumption 4. $\omega > \underline{\phi}$.

We have considered the developer's optimal privacy setting assuming that their choice would be compatible with any minimum privacy requirements that could be placed on them by the platform. However, if that platform minimum standard was binding, constraining their privacy choice, we would need to modify the best response to:

$$\phi_a = \begin{cases} \omega - \underline{\phi} & \text{if } \omega - \underline{\phi} \geq \phi_{min}, \\ \phi_{min} & \text{if } \omega - \underline{\phi} < \phi_{min}. \end{cases} \tag{3}$$

Substituting the optimal solution and the best response into the developer profit function, we obtain the individual rationality constraint for a developer to be active in the market.

Lemma 3. *Developers choose to actively participate on the platform so long that:*

$$\theta_a \leq \begin{cases} \dfrac{(\omega - \underline{\phi})^2}{2} & \textit{if } \omega - \underline{\phi} \geq \phi_{min}, \\ \dfrac{(\omega + \phi_{min} - \underline{\phi})^2}{4} - \dfrac{(\phi_{min})^2}{2} & \textit{if } \omega - \underline{\phi} < \phi_{min}. \end{cases} \tag{4}$$

With this definition, we will also describe the marginal participating developer's type $\tilde{\theta}$ to be:

$$\tilde{\theta} = \begin{cases} \dfrac{(\omega - \underline{\phi})^2}{2} & \text{if } \omega - \underline{\phi} \geq \phi_{min}, \\ \dfrac{(\omega + \phi_{min} - \underline{\phi})^2}{4} - \dfrac{(\phi_{min})^2}{2} & \text{if } \omega - \underline{\phi} < \phi_{min}. \end{cases} \tag{5}$$

Platform A platform imposing a minimum required level of privacy to developers, where $\omega - \underline{\phi} < \phi_{min}$, leads to a platform profit function defined as:

$$\Pi = \int_{\underline{\phi}}^{\frac{\omega + \phi_{min} + \underline{\phi}}{2}} f(\phi)\, d\phi \int_{\underline{\theta}}^{\frac{(\omega + \phi_{min} - \underline{\phi})^2}{4} - \frac{(\phi_{min})^2}{2}} g(\theta)\, d\theta. \tag{6}$$

Lemma 4. $\int_{\underline{\phi}}^{\frac{\omega + \phi_{min} + \underline{\phi}}{2}} f(\phi)\, d\phi$ *captures the mass of users who participate on the platform and install apps.*

Lemma 5. $\int_{\underline{\theta}}^{\frac{(\omega + \phi_{min} - \underline{\phi})^2}{4} - \frac{(\phi_{min})^2}{2}} g(\theta)\, d\theta$ *captures the mass of active developers (who are developing apps).*

Conceptually, this specification interprets platform profit as the gains derived from user traffic, i.e the data the platform extracts and utilises from active users (in the sense of users installing those apps that have been developed).

Acknowledging the fact that the platform itself is always harvesting data, this specification highlights that the data obtained from users will differ per application. As an example, Google Maps (a mapping/GPS guiding application) provides user location data to the platform, while a banking application provides data about the user's financial transactions. The platform gathers all this data and utilises it for their own benefit, thus translating into profit. This is where the outcome of a single user i counting as multiple users is highlighted, as this aggregates all of the applications that i uses, and thus all the data that is extracted.

Evaluating this profit function and taking the partial derivative with respect to the minimum privacy standard ϕ_{min} gives the optimal minimum privacy standard as a function of the model's core variables.

Proposition 1. *The platform's best response function for a binding minimum privacy standard ϕ_{min} corresponds to*
$\phi_{min} = \frac{1}{3}(\omega - \underline{\phi}) + \sqrt{\frac{1}{9}(\omega + \underline{\phi})^2 + (\omega - \underline{\phi})^2 - \frac{4}{3}\underline{\theta}}.$

Furthermore, the optimal level for the minimum privacy standard need to be non-negative. Additionally, the root term must also be positive. We limit this analysis only to the reals. This places a constraint on the size of $\underline{\theta}$, the lowest possible development cost:

Assumption 5. *The lowest possible development cost $\underline{\theta}$ is bound such that* $\underline{\theta} \leq \frac{10\omega^2 - 17\omega\underline{\phi} + 10\underline{\phi}^2}{12}$.

Below we summarise how the platform's optimal minimum privacy standard changes with respect to $\omega, \underline{\phi}$, and $\underline{\theta}$.

Partial derivative	Sign
$\frac{\partial}{\partial \omega}$	$+$
$\frac{\partial}{\partial \underline{\phi}}$	$-$
$\frac{\partial}{\partial \underline{\theta}}$	$-$

These signs make logical sense. With increases in the lowest possible level of privacy $\underline{\phi}$ and the lowest possible development cost $\underline{\theta}$, it makes sense that the minimum privacy standard would decrease, in order to ensure that standards are not set too high to restrict developer participation. Conversely, an increase in ω implies a larger mass of both users and developers participating, which allows the platform to institute higher standards.

So far we have assumed that developers would be constrained by the platform's privacy standard, i.e. $\omega - \underline{\phi} < \phi_{min}$. We can now formalise this by placing an additional bound for the relative size of $\underline{\theta}$.

Proposition 2. *The platform's minimum privacy standard will be binding when* $\underline{\theta} < \frac{6\omega^2 - 8\omega\underline{\phi} + 6\underline{\phi}^2}{12}$.

Furthermore, when considering the scenario in which the optimal developer privacy level is greater than or equal to the minimum standard (the unconstrained outcome), the crucial juncture at which this deviates away from the previous analysis is in the formulation of the platform profit function. In this instance, the platform would see a profit function of:

$$\Pi = \int_{\underline{\phi}}^{\omega} f(\phi)\, d\phi \int_{\underline{\theta}}^{\frac{(\omega-\underline{\phi})^2}{2}} g(\theta_a)\, d\theta_a. \tag{7}$$

Lemma 6. $\int_{\underline{\phi}}^{\omega} f(\phi)\, d\phi$ *captures the mass of users who participate on the platform and install apps when developers are not restricted by the platform minimum privacy standard.*

Lemma 7. $\int_{\underline{\theta}}^{\frac{(\omega-\underline{\phi})^2}{2}} g(\theta)\, d\theta$ *captures the mass of active developers (who are developing apps) when their optimal privacy level* ϕ_a *exceeds the platform minimum standard.*

From here, the platform does not need to take any actions. There is no optimal solution for ϕ_{min} as it would not be relevant.

2.2 Comparative Statics

We perform comparative statics, observing how the size of ω, $\underline{\phi}$, and $\underline{\theta}$ will affect agent actions. We also analyse whether the platform imposing a minimum standard is beneficial for social welfare.

We consider six outcomes in our comparative statics: user mass, aggregate net user utility, aggregate gross user utility, developer mass, developer profit, and platform profit.

While the other outcomes have been previously defined in our analysis, we have yet to formalise the aggregate user utilities. Aggregate net user utility will be defined as the total utility seen by all users.

Comment 2. *U_{net} is defined as the aggregate net user utility for the participating mass of users*

$$U_{net} = \int_{\underline{\phi}}^{\tilde{\phi}_i} u_{i,a}\, f(\phi)\; d\phi \int_{\underline{\theta}}^{\tilde{\theta}_a} g(\theta)\; d\theta. \tag{8}$$

Note that by substituting in the marginal participating agent condition as well as the best response functions for p_a, U can be written in generalised form:

$$U_{net} = \frac{(\omega + \phi_a - 3\underline{\phi})^2}{8} \left(\frac{(\omega + \phi_a - \underline{\phi})^2}{4} - \frac{\phi_a^2}{2} - \underline{\theta} \right).$$

This is simply the aggregate net utility seen by all participating users across all applications (and thus, on the platform). Each of the integrals define the appropriate participating mass (users and developers respectively) with accordance to the marginal participating condition $\widetilde{\phi}$ and $\widetilde{\theta}$.

We also calculate the aggregate gross user utility, which accounts for the utility received by the entire mass of participating consumers while disregarding the price to be paid to install these applications (p_a).

Comment 3. *U_{gross} is defined as the aggregate gross user utility for the participating mass of users*

$$U_{gross} = \int_{\underline{\phi}}^{\widetilde{\phi}} (\omega + \phi_a - \phi_i)\, f(\phi)\; d\phi \int_{\underline{\theta}}^{\widetilde{\theta}} g(\theta)\; d\theta. \tag{9}$$

Compared to the aggregate net utility, we replace $u_{i,a}$ with $(\omega + \phi_a - \phi_i)$ to calculate the gross user utility, removing the price paid to developers (p_a).

Note that by substituting in the marginal participating agent condition as well as the best response functions for p_a, U can be written in generalised form:

$$U_{gross} = \frac{3(\omega + \phi_a - \underline{\phi})^2}{8} \left(\frac{(\omega + \phi_a - \underline{\phi})^2}{4} - \frac{\phi_a^2}{2} - \underline{\theta} \right).$$

We hold our previous assumptions (1 and 2) to be true, and take the partial derivative for each outcome.

Partial derivative	User mass	Aggregate net utility	Aggregate gross utility
$\frac{\partial}{\partial \omega}$	+	+	+
$\frac{\partial}{\partial \underline{\phi}}$	−	−	−
$\frac{\partial}{\partial \underline{\theta}}$	−	−	−

Partial derivative	Developer mass	Developer profit	Platform profit
$\frac{\partial}{\partial \omega}$	+	+	+
$\frac{\partial}{\partial \underline{\phi}}$	−	−	−
$\frac{\partial}{\partial \underline{\theta}}$	−	−	−

We get consistent outcomes, with an increase in ω resulting in better payoffs for all, driven by the increase in user utility (this leads to a higher user mass and developer profit, which leads to a larger developer mass, and thus, a higher platform profit). We see the inverse is true with the two constraint variables, $\underline{\phi}$ and $\underline{\theta}$. Increasing $\underline{\phi}$ will lead to users having stricter privacy standards, and thus, a smaller user mass. This flows on to decrease the payoffs for all agents. An increase in $\underline{\theta}$ will affect developers, reducing the participating mass. This also affects users, and platform profit.

2.3 Social optimum for privacy settings

We now consider the welfare of the entire model. We define the aggregate welfare W as:

$$W = \int_{\underline{\phi}}^{\widetilde{\phi}} (\omega + \phi_a - \phi_i)\, f(\phi)\, d\phi \int_{\underline{\theta}}^{\widetilde{\theta}} g(\theta)\, d\theta - \left(\theta_a + \frac{\phi_a}{2}\right) \int_{\underline{\theta}}^{\widetilde{\theta}} g(\theta)\, d\theta. \tag{10}$$

The first part of this expression denotes the gross aggregate utility experienced by users. Thus, the benefit that they receive from using all active applications available in the market, disregarding the price (p_a) that they would pay. Fundamentally p_a is the revenue (and thus welfare) seen by developers for each user on an application, therefore, we know that the net welfare is not affected by p_a and rather than using the net aggregate utility, we simply use the gross aggregate utility to simplify.

The second part of this expression denotes the cost borne by the developers with respect to participation and implementation of their chosen privacy level. This accounts for all considered costs in the model.

By taking the partial derivative of W with respect to ϕ_a, we can derive the socially optimal privacy level ϕ^W:

Proposition 3. *The socially optimal privacy level is equal to:*

$$\phi^W = \frac{1}{72}\left(18 + 18\Delta + \sqrt[3]{\frac{1944}{\mu}}(10\Delta - 23\Delta^2 + 24\underline{\theta} - 3) - \sqrt[3]{648\mu}\right). \quad (11)$$

In which we define Δ and μ as:

$$\Delta = \omega - \underline{\phi},$$

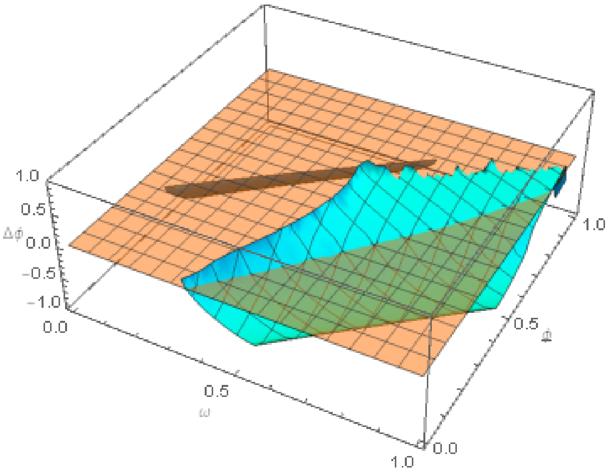
$$\mu = (9 - 27\Delta)(9\Delta^2 + 2\Delta - 20\underline{\theta} - 1) + \frac{1}{3}\sqrt{(30\Delta - 69\Delta^2 + 72\underline{\theta} - 9)^3 + 729(1 - 3\Delta)^2(1 - 2\Delta - 9\Delta^2 + 20\underline{\theta})^2}.$$

When comparing the socially optimal privacy setting ϕ^W along with ϕ_{min} and ϕ_a (the profit maximising privacy setting for developers), outcomes differ based on the relative sizes of ω, $\underline{\phi}$, and $\underline{\theta}$. Comparisons between these settings are difficult to identify due to the variables being mostly unconstrained, however, when analysing the functions from a graphical perspective, we see scenarios in which $\phi^W > \phi_{min}$ and vice versa, thus we posit the following corollary.

Corollary 1. *There exists a set of values for $\underline{\theta}$ while $\omega > \underline{\phi}$, in which both positive real values for ϕ^W and ϕ_{min} exist. Furthermore, at least one set of values will hold $\phi_{min} > \phi^W$, and at least one other set will hold $\phi^W > \phi_{min}$.*

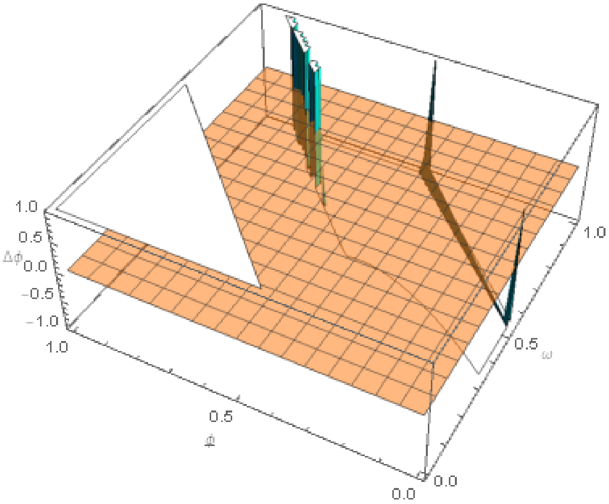
We present an image below that demonstrates the validity of this corollary, as the difference between ϕ^W and ϕ_{min} has values both above and below the zero-level-orange-plane.

Figure 1: The difference between ϕ^W and ϕ_{min} ($\underline{\theta} = 0.1$ and $\omega, \underline{\phi} \in (0,1)$



A crucial factor in determining whether the socially optimal privacy level is higher or lower than the profit maximising platform privacy standard is ω (the utility that users experience from consuming an application).

Figure 2: The partial derivative of the difference between ϕ^W and ϕ_{min} with respect to ω ($\underline{\theta} = 0.1$ and $\omega, \underline{\phi} \in (0,1)$



Due to the complexity of the function, it is difficult to directly state the effect that changing ω will have. However, considering the effect that ω has with other outcomes (such as user and developer mass, alongside developer and platform profit), an increasing ω would imply an overall increase in aggregate welfare.

3 Conclusion

We provide a model in which we demonstrate how agents behave with respect to privacy settings on a digital platform. We find best response functions for developer pricing, along with optimal values for the developer level profit maximising privacy setting, the platform level profit maximising privacy setting (through calculating their desired minimum level of privacy) along with the full game welfare maximising privacy setting.

We discover that agent outcomes are affected consistently by the relevant exogenous variables, and also find that the relevant privacy choice will depend on the relative size of these variables. We are able to conclude that the platform's minimum privacy standard may not always be binding, and that the higher the minimum development cost is for developers, the more likely it is that their optimal privacy setting is greater than the platform standard.

We also conclude the platform minimum privacy standard is not always welfare maximising. We discover a welfare maximising privacy standard, and have identified scenarios in which it is both greater and less than the platform privacy standard. Due to the complexity in calculations, we posit a corollary in which we predict that the size of the lowest possible development cost will affect whether the welfare maximising privacy setting is above or below the platform standard.

References

Acemoglu, D., Makhdoumi, A., Malekian, A., and Ozdaglar, A. (2019). Too much data: Prices and inefficiencies in data markets. Technical report, National Bureau of Economic Research.

Acquisti, A., Taylor, C., and Wagman, L. (2016). The economics of privacy. *Journal of economic Literature*, 54(2):442–92.

Apple (2021). Privacy.

Bergemann, D. and Bonatti, A. (2015). Selling cookies. *American Economic Journal: Microeconomics*, 7(3):259–94.

Bergemann, D., Bonatti, A., and Gan, T. (2020). The economics of social data.

Board, S. and Lu, J. (2018). Competitive information disclosure in search markets. *Journal of Political Economy*, 126(5):1965–2010.

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

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

Fainmesser, I. P., Galeotti, A., and Momot, R. (2019). Digital privacy. *HEC Paris Research Paper No. MOSI-2019-1351.*

Hagiu, A. and Jullien, B. (2011). Why do intermediaries divert search? *The RAND Journal of Economics*, 42(2):337–362.

Hirshleifer, J. (1980). Privacy: Its origin, function, and future. *The Journal of Legal Studies*, 9(4):649–664.

Ichihashi, S. (2020). Dynamic privacy choices. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 539–540.

Noam, E. M. (1997). Privacy and self-regulation: Markets for electronic privacy. *Privacy and Self-Regulation in the Information Age*, pages 21–33.

Parikh, P. (2021). Signal hits 50 million installs on play store amid whatsapp privacy concerns.

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

Rosenberg, S. (2019). Here's what apple, facebook and google really mean when they talk about "privacy".

Stigler, G. J. (1961). The economics of information. *Journal of political economy*, 69(3):213–225.

Stigler, G. J. (1980). An introduction to privacy in economics and politics. *The Journal of Legal Studies*, 9(4):623–644.

Varian, H. (2002). Economic aspects of personal privacy. cyber policy and economics in an internet age, 127-137.

WhatsApp (2020). Two billion users–connecting the world privately.

WhatsApp (2021). Privacy policy.

Yang, K. H. (2020). Selling consumer data for profit: Optimal market-segmentation design and its consequences.

A Proofs

Lemma 1. *The developer's best response pricing p_a is $\frac{\omega+\phi_a-\underline{\phi}}{2}$.*

Proof. Taking the partial derivative of the developer profit function (2) with respect to p_a gives $\omega - 2p_a + \phi_a - \underline{\phi}$. Setting this to equal 0 and rearranging provides the best response pricing of $\frac{\omega+\phi_a-\underline{\phi}}{2}$. ■

Lemma 2. *The developer's optimal privacy setting ϕ_a given its best response pricing will be $\omega - \underline{\phi}$.*

Proof. Substituting in the best response pricing into the developer profit function, then taking the partial derivative with respect to ϕ_a will give $\omega - \phi_a - \underline{\phi}$. Setting this equal to 0 and rearranging with respect to ϕ_a gives $\omega - \underline{\phi}$. ■

Lemma 3. *Developers will choose to participate on the platform based on the relationship between the best response privacy setting and the platform's privacy standard. This will lead to two possible participation conditions:*

$$\theta_a \leq \begin{cases} \dfrac{(\omega-\underline{\phi})^2}{2} & \textit{if } \omega - \underline{\phi} \geq \phi_{min}, \\[2ex] \dfrac{(\omega+\phi_{min}-\underline{\phi})^2}{4} - \dfrac{(\phi_{min})^2}{2} & \textit{if } \omega - \underline{\phi} < \phi_{min}. \end{cases}$$

Proof. We begin by substituting in the derived best response functions into the original developer profit function (2). This gives: $\pi_a = (\omega - \underline{\phi})^2 - \theta_a - \left(\frac{(\omega-\underline{\phi})^2}{2}\right)$. Setting profit π_a equal to 0 allows us to rearrange for the marginal type of developer who would still participate, leading to the condition of $\theta_a \leq \frac{(\omega-\underline{\phi})^2}{2}$.

This only considers the scenario in which the best response privacy setting ϕ_a is greater than the minimum privacy standard ϕ_{min}. To account for this scenario, we substitute in the developer best response pricing as done previously, but also substitute in ϕ_{min} rather than the best response privacy setting ϕ_a. This gives $\pi_a = \left(\frac{\omega+\phi_{min}-\underline{\phi}}{2}\right)^2 - \theta_a - \frac{\phi_{min}^2}{2}$. Setting π_a equal to 0 and rearranging gives the participation condition of $\theta_a \leq \frac{(\omega+\phi_{min}-\underline{\phi})^2}{4} - \frac{(\phi_{min})^2}{2}$ while $\phi_a < \phi_{min}$. ■

Lemma 4. $\displaystyle\int_{\underline{\phi}}^{\frac{\omega+\phi_{min}+\underline{\phi}}{2}} f(\phi)\, d\phi$ *captures the mass of users who participate on the platform and install apps.*

Proof. The marginal user for application a will have a desired privacy level $\widetilde{\phi}$ derived from $\phi_i = \omega - p_a + \phi_a$. Substituting in the best response function for a's price $p_a = \frac{\omega+\phi_a-\underline{\phi}}{2}$ along with the optimal privacy setting (under a platform imposed minimum privacy standard) $\phi_a = \phi_{min}$ gives $\widetilde{\phi} = \frac{\omega+\phi_{min}+\underline{\phi}}{2}$, giving the upper bound of participating users. ■

Lemma 5. $\int_{\underline{\theta}}^{\frac{(\omega+\phi_{min}-\underline{\phi})^2}{4}-\frac{(\phi_{min})^2}{2}} g(\theta)\, d\theta$ *captures the mass of active developers (who are developing apps).*

Proof. We have identified the marginal developer under a platform imposed minimum privacy setting as having a development cost that fulfils $\theta_a = \frac{(\omega+\phi_{min}-\underline{\phi})^2}{4} - \frac{(\phi_{min})^2}{2}$ and thus, we set this as the upper bound for the participating developers and denote this as $\widetilde{\theta}$. ■

Proposition 1. *The platform's best response function for the minimum privacy standard ϕ_{min} given the developer best response functions is defined as:* $\phi_{min} = \frac{1}{3}(\omega-\underline{\phi}) + \sqrt{\frac{1}{9}(\omega+\underline{\phi})^2 + (\omega-\underline{\phi})^2 - \frac{4}{3}\underline{\theta}}$.

Proof. First, we evaluate the platform profit function and obtain:

$$\Pi = \left(\frac{\omega+\phi_{min}-\underline{\phi}}{2}\right)\left(\left(\frac{(\omega+\phi_{min}-\underline{\phi})}{2}\right)^2 - \frac{(\phi_{min})^2}{2} - \underline{\theta}\right).$$

Taking the partial derivative of this function with respect to ϕ_{min} gives

$$\frac{3(\omega^2+\underline{\phi}^2-\phi_{min}^2)}{8} + \frac{\phi_{min}(\omega-\underline{\phi})}{4} - \frac{3\omega\underline{\phi}}{4} - \frac{\underline{\theta}}{2}.$$

We set this to equal 0 to find the best response ϕ_{min}. An immediate outcome of this is that there will be two possible solutions for ϕ_{min}. Continuing to solve, we obtain the two expressions:

$$\phi_{min} = \begin{cases} \frac{1}{3}(\omega-\underline{\phi}) + \sqrt{\frac{1}{9}(\omega+\underline{\phi})^2 + (\omega-\underline{\phi})^2 - \frac{4}{3}\underline{\theta}}, \\ \frac{1}{3}(\omega-\underline{\phi}) - \sqrt{\frac{1}{9}(\omega+\underline{\phi})^2 + (\omega-\underline{\phi})^2 - \frac{4}{3}\underline{\theta}}. \end{cases}$$

We first find the constraint that ensures that the root expression is positive. This is highlighted in assumption 2 with $\underline{\theta} \leq \frac{10\omega^2-17\omega\underline{\phi}+10\underline{\phi}^2}{12}$ being the condition that needs to be held for us to remain in the reals.

Given that to be true, we then know that

$$\frac{1}{3}(\omega - \underline{\phi}) + \sqrt{\frac{1}{9}(\omega + \underline{\phi})^2 + (\omega - \underline{\phi})^2 - \frac{4}{3}\underline{\theta}},$$

will indeed be the profit maximising ϕ_{min}, as it will be larger than the alternative expression, ensuring that it is the profit maximising ϕ_{min}, not minimising. ■

Proposition 2. *The platform's minimum privacy standard will be binding when* $\underline{\theta} < \frac{6\omega^2 - 8\omega\underline{\phi} + 6\underline{\phi}^2}{12}$.

Proof. We first establish that $\phi_{min} > \phi_a$ must be true for the platform minimum standard to be in effect. When substituting in the optimised functions, we find this inequality:

$$\frac{1}{3}(\omega - \underline{\phi}) + \sqrt{\frac{1}{9}(\omega + \underline{\phi})^2 + (\omega - \underline{\phi})^2 - \frac{4}{3}\underline{\theta}} > \omega - \underline{\phi}.$$

We then evaluate this inequality to obtain the condition for $\underline{\theta}$:

$$\frac{6\omega^2 - 8\omega\underline{\phi} + 6\underline{\phi}^2}{12} > \underline{\theta}.$$

Thus, we find the constraint on $\underline{\theta}$ for when the optimal platform minimum privacy standard will hold from it being larger than the optimal developer privacy setting. ■

Lemma 6. $\int_{\underline{\phi}}^{\omega} f(\phi)\, d\phi$ *captures the mass of users who participate on the platform and install apps when developers are not restricted by the platform minimum privacy standard.*

Proof. Similar to the constrained scenario in which the minimum privacy standard is active, the marginal user for application a will have a desired privacy level of $\widetilde{\phi} = \omega - p_a + \phi_a$. The developer's best response pricing $p_a = \frac{\omega + \phi_a - \underline{\phi}}{2}$ remains the same, however privacy setting is now $\phi_a = \omega - \underline{\phi}$. Thus, when substituting in these values find the marginal user's privacy demand, we obtain $\widetilde{\phi} = \omega$. ■

Lemma 7. $\int_{\underline{\theta}}^{\frac{(\omega - \underline{\phi})^2}{2}} g(\theta)\, d\theta$ *captures the mass of active developers (who are developing apps) when their optimal privacy level* ϕ_a *exceeds the platform minimum standard.*

Proof. We have identified the marginal developer under a unconstrained environment as having a development cost that fulfils $\frac{(\omega-\underline{\phi})^2}{2}$ and thus, we set this as the upper bound for the unconstrained environment. ■

Proposition 3. *The socially optimal privacy level is equal to:*

$$\phi^W = \frac{1}{72}\left(18 + 18\Delta + \sqrt[3]{\frac{1944}{\mu}}(10\Delta - 23\Delta^2 + 24\underline{\theta} - 3) - \sqrt[3]{648\mu}\right).$$

Proof. We first evaluate the aggregate welfare function W with respect to the participation conditions of the marginal user and developer, along with the best response pricing function for p_a:

$$\begin{aligned} W = {} & \frac{3(\omega - \underline{\phi} + \phi)^2}{8}\left(\frac{(\omega - \underline{\phi} + \phi)^2}{4} - \frac{\phi^2}{2} - \underline{\theta}\right) \\ & - \left(\frac{(\omega - \underline{\phi} + \phi)^2}{4} - \frac{\phi^2}{2} - \underline{\theta}\right)\left(\frac{(\omega - \underline{\phi} + \phi)^2}{8} - \frac{\phi^2}{4} + \frac{\underline{\theta}}{2} + \frac{\phi}{2}\right). \end{aligned}$$

We utilise the average values for participating agents when it comes to defining ϕ_i and θ_a, due to both being drawn from uniform distribution functions. Furthermore, we replace ϕ_a, the developer's privacy setting, with ϕ.

We now take the partial derivative of W with respect to ϕ:

$$\begin{aligned} \frac{\partial W}{\partial \phi} = {} & \frac{1}{8}(-4\phi^3 + 4\underline{\theta} - 6\underline{\theta}\phi - 6(\omega - \underline{\phi})\underline{\theta} - 4(\omega - \underline{\phi})\phi \\ & + 5(\omega - \underline{\phi})^2\phi - (\omega - \underline{\phi})^2 + 2(\omega - \underline{\phi})^3 + 3\phi^2 + 3(\omega - \underline{\phi})\phi^2). \end{aligned}$$

This simplifies into a cubic function:

$$\begin{aligned} \frac{\partial W}{\partial \phi} = {} & -4\phi^3 + \big(3 + 3(\omega - \underline{\phi})\big)\,\phi^2 + \big(5(\omega - \underline{\phi})^2 - 4(\omega - \underline{\phi}) - 6\underline{\theta}\big)\,\phi \\ & + \big[2(\omega - \underline{\phi})^3 - (\omega - \underline{\phi})^2 - 6(\omega - \underline{\phi})\underline{\theta} + 4\underline{\theta}\big]. \end{aligned}$$

We now set this function to equal 0 in order to solve for the optimal value of ϕ. When calculating, we find three possible solutions, two of which are in the complex plane. Remaining in the reals, we derive the only viable solution for ϕ. Note that this intrinsically implies that the solution will be positive. To simplify our expression, we abbreviate two terms:

$$\Delta = \omega - \underline{\phi},$$

$$\mu = (9 - 27\Delta)(9\Delta^2 + 2\Delta - 20\underline{\theta} - 1) + \\ \frac{1}{3}\sqrt{(30\Delta - 69\Delta^2 + 72\underline{\theta} - 9)^3 + 729(1 - 3\Delta)^2(1 - 2\Delta - 9\Delta^2 + 20\underline{\theta})^2}.$$

Thus we define ϕ:

$$\phi^W = \frac{1}{72}\left(18 + 18\Delta + \sqrt[3]{\frac{1944}{\mu}}(10\Delta - 23\Delta^2 + 24\underline{\theta} - 3) - \sqrt[3]{648\mu}\right).$$

■