UNIVERSITY OF MILAN
AND
UNIVERSITY OF PAVIA

DOCTORAL THESIS

Essay on Advertising in Digital Platforms and their Taxation

Author:
Nicola CHIAROMONTE

Supervisor:
Prof. Luca MICHELETTO

A thesis submitted in fulfilment of the requirements for the degree of Economics – Doctor of Philosophy

Academic Year 2019/20

Contents

Chapter 1

Introduction: digitalization of economy and related issues

1.1 The era of digitalization

Year by year we are witnesses of fast transformation of our society, of our habits, of our life thanks to new technologies and new ways of interaction. On one side it is impossible not to state that a great benefit for us comes from this digital revolution, on the other side new challenges come for us in terms of deregulation and obsolescence of the current rules system. It is not always easy to face new situations linked to innovation by using standard tools. Moreover the intangibility of these new means and the widening of the range of their action contribute to blur national borders and call for a new international law system.

We should consider in this context the interest for this phenomenon showed by several governments and their demand for regulation addressed to international institutions, such as the European Commission (EC) or the OECD.

Before wrestling, however, with these issues and how they have been so far faced by single governments and supranational organizations, we ought to outline the main questions arisen by digitalization.

The connectivity between users and devices generates a huge data exchange. General and personal data is collected by businesses and governments and used to develop and offer services in order to answer old and new needs. It is the good side of the digital revolution, but other issues, less clear and not always positive, come when innovation advances. Indeed, new technologies raise concerns linked to transparency, control and security of personal information, and related legal jurisdiction. The novelty and intangibility of this economy allow to create businesses not subjected to any national laws. This aspect constitutes an additional benefit with respect to traditional "brick and mortar" businesses, especially in terms of fiscal gains.

1.1.1 Main features

Digital markets present several features. They have doubtlessly allowed businesses to locate some stages of their production chain across different countries, exploiting local advantages, economic and not. Furthermore, digital markets have made possible to address customers worldwide without actual physical presence, or as the EC says without a “permanent establishment”.

Another key element consists of the growing use of IP assets, like software and algorithms, as base for building platforms and websites. Their relative weight in the chain value has been growing year by year.

In addition, digital enterprises are commonly characterized by data exploitation, user participation, network effects and the provision of user-generated content. The degree of active and passive user participation can be seen as a form of “free” work, as stated by Colin and Collin (2013), when it contributes to create value for businesses.

These markets are characterized by direct and/or indirect network effects. Utility from the consumption of a specific good or service can depend on the number of other end-users consuming the same good or service, for instance trip or food review sites, or from the interaction between two groups, such as the users and the advertisers of a social network.

High fixed costs, associated for instance to an algorithm development, and very low variable costs allow digital platforms to exploit economies of scale. This feature, in combination with huge switching costs, due do lock-in effects, gives them a great market power. In fact, the biggest actors of digital platforms operate in a monopoly, or at least oligopoly, by offering a specific and distinct service to users. As a result, digital markets are often not competitive, since single firms can influence market prices, reducing the consumer surplus of their users.

A representative example of this world can be a social platform supported by advertising revenue. A social network is a multi-sided platform that collects users' data and provides advertising services. It offers specific services to users in exchange of a payment of a fee, which in some cases can be also zero. Moreover, it collects users' data, not only linked to platform usage, but also personal information about habits, tastes, and so on. On the other side of the market, the platform enables firms to reach a target audience though advertising, and sometimes to sell directly their products in an effective and efficient manner. Revenue may come from one of the two sides of the market, or from both. However, digital businesses exploit personal data for users' experience improvement with the aim of selling more and better targeted advertising space to third parties, which may in turn increase their sales.
To conclude, the specific business model presented here is only one of many possibilities to make profits through digitalization. It allows us to focus mainly on two problems: the privacy protection of users and the unfair tax advantage with respect to traditional businesses.

1.1.2 The data protection issue

Digital giants have built up profitable empires by observing their consumers' online behaviour and collecting their data. The importance of "big data" in new forms of business raises two separate issues. First of all, as highlighted in the French Report by Colin and Collin (2013), data are a relevant input in the value chain for many digital platforms, and users voluntarily, but sometimes not consciously, upload them without an exchange of money. The absence of any financial transaction, but a de facto barter transaction between data and services, makes the fiscal authorities unable to properly tax the benefit coming from their own citizens, creating a distortion and an unfair advantage with respect to other sectors and businesses.
Moreover, a second problem is linked to the "fairness" of the exchange between

users and platforms, in particular whether the users receive a fair share of the surplus. In fact, the data exploitation, while providing a valuable service to users by improving platform's experience and targeted ads, instead of advertising junk, also involves a cost in terms of privacy loss. This last aspect is more relevant when platforms engage unknown intermediaries in the resale of data through opaque arrangements, which finally results in a loss of control on the dissemination of personal data to third parties, as Facebook–Cambridge Analytica data scandal[1] has shown, for instance.

The privacy cost associated with the use of new technologies has been tackled mainly by the European Commission, and it is still an object of concern. Indeed, the EU General Data Protection Regulation (GDPR) is the most important change in data privacy regulation done so far. The aim of the GDPR directive is to protect all EU citizens from personal data breaches in today's data-driven world.

Nevertheless, solving this problem will require additional efforts and it will not be easy to separate the blurring edge between benefits and costs of personal data.

1.1.3 The fiscal issue

A different problem linked to digitalization is that the current international corporate tax rules are not appropriate for the realities of the modern global economy and fail to capture those businesses that make profits from digital services in a country without being physically present there. Current tax system, moreover, fails to recognise how profits are created in the digital world, in particular the role that users play in generating value for digital companies. As a result, there is a "mismatch" between where value is created and where taxes are paid, when they are.

According to European Commission (2018), "in the digital economy, value is often created from a combination of algorithms, user data, sales functions and knowledge.

[1]See for instance the article on The Economist (2018).

For example, a user contributes to value creation by sharing his/her preferences (e.g. liking a page) on a social media forum. This data will later be used and monetised for targeted advertising. The profits are not necessarily taxed in the country of the user (and viewer of the advert), but rather in the country where the advertising algorithms has been developed, for example. This means that the user contribution to the profits is not taken into account when the company is taxed."
The point here is that digital companies benefit from the same advantages and infrastructure as traditional businesses, such as high speed internet, roads, a stable legal system and others. The problem is that often, they do not contribute by paying their fair share of taxes in the countries they do business in.
To answer this necessity, so far, unilateral actions have been undertaken by single governments, ending up with complex and not so clear law systems, causing double taxation problems. In order to overcome the fear of facing as many fiscal authorities as the markets that digital businesses serve, and to avoid excessive tax compliance burdens, which could be a stop for innovation, there is today the need for international rules. Under this reading key, the European Commission and the OECD are seeking to find a shared solution, capable to cover as many cases as possible.

1.1.4 The European Commission proposal

On 21 March 2018, the European Commission proposed new rules to ensure that digital business activities in the EU are taxed in a fair and growth-friendly way.
The Commission has made two legislative proposals:

- The first initiative aims to reform corporate tax rules, so that profits are registered and taxed where businesses have significant interaction with users through digital channels. It is the Commission's preferred long-term solution.

- The second proposal responds to calls from several Member States for an interim tax which covers the main digital activities that currently escape taxation altogether in the EU.

The first proposal would enable Member States to tax profits that are generated in their territory, even if a company does not have a physical presence there. The new rules would ensure that online businesses contribute to public finances at the same level as traditional "brick-and-mortar" companies.

A digital platform will be deemed to have a taxable "digital presence" in a Member State if it fulfils one of these three criteria: more than €7 million in annual revenues, or more than 100.000 users, or over 3.000 business contracts for digital services in a Member State in a taxable year.

Moreover, the new rules will also change the way profits are allocated to Member States to secure a real link between where digital profits are made and where they are taxed. The measure could eventually be integrated into the scope of the Common Consolidated Corporate Tax Base (CCCTB).

The second proposal suggests an interim tax aimed at generating immediate revenues for Member States from those activities which are currently not effectively taxed. Moreover, it would avoid unilateral measures to tax digital activities in certain Member States which could lead to a patchwork of national responses that would be damaging for the EU Single Market.

Until the comprehensive reform has been implemented, the tax would apply to revenues created from activities where users play a major role in value creation and which are the hardest to capture with current tax rules. Examples are sales of online advertising space, financial transactions on a platform when this can facilitate the exchange of goods and services between users, sales of data generated from user-provided information.

According to the proposal, tax revenues would be collected by the Member States

where the users are located, and will only apply to big companies with total annual worldwide revenues of €750 million and EU revenues of €50 million, whereas smaller start-ups and scale-up businesses will remain unburdened.
Though the legislative proposals have not been applied so far, the EU actively contributes to the global discussions on digital taxation within the G20/OECD, and pushes for ambitious international solutions.

1.1.5 The OECD proposal

Tax challenges of the digitalization of the economy have been identified as one of the main focus of the Base Erosion and Profit Shifting (BEPS) Action Plan by the Organization for Economic Co-operation and Development (OECD), in particular the impossibility to ring-fence the digital economy as observed by the 2015 BEPS Action 1 Report.
A whole and deep analysis of this phenomenon has been the object of an Interim Report, delivered in March 2018 by the Task Force on the Digital Economy (TFDE), entitled "Tax Challenges Arising from digitalization - Interim Report 2018". It provides a study of value creation across new and changing business models in the context of digitalization and it identifies three frequently observed characteristics: scale without mass has a direct consequence on the country's taxing right; a heavy reliance on intangible assets exploits uncertainties and opportunities to relocate income; data and user participation play a crucial role in the chain value, but not in profit allocation rules.
A common conclusive agreement has not yet been reached, though all the members agreed to continue working together to a final report in 2020. Nevertheless, in the Policy Note "Addressing the Tax Challenges of the digitalization of the Economy, approved on 23 January 2019, the Inclusive Framework proposed two pillars as the basis for consensus:

- Pillar One tackles the allocation of taxing rights between different jurisdictions;
- Pillar Two focuses on developing a rule to “tax back” when the fiscal subject faces a too low level of effective taxation in the jurisdictions having their primary taxing rights;

According to the former, the Inclusive Framework is looking for changes to the permanent establishment rule, giving space to new concepts, such as “significant economic presence” or “significant digital presence”.
With regard to the latter, the global anti-base erosion (GloBE) proposal would address the risk of profit shifting and the risk of un-coordinated, unilateral action, in absence of a multilateral agreement, which would finally harm all countries, large and small, developed and developing. It would consist of two inter-related rules: an income inclusion rule to allow taxation when the effective tax rate is below a minimum rate (this seems very similar to the actual CFC rules); and a tax on base eroding payments which would deny a deduction or impose a withholding tax unless these payments are at or above a minimum rate.
Although the OECD is currently working on an international solution for digital taxation, a number of countries have recently taken, or are going to take, unilateral measures to implement a digital services tax. We are showing in the next paragraph who they are and how they are facing the digitalization issues.

1.1.6 The unilateral national answers

With no consensus on taxation of the digital economy, some countries have resorted to unilateral measures. Such measures are broadly of four kinds:

1. alternative applications of the permanent establishment threshold (such as “significant presence” tests or “virtual” permanent establishments);

2. withholding taxes (in particular for industries such as advertising, broader definitions of royalties);
3. "equalization" levies on internet advertising and digital services taxes;
4. specific regimes to deal with large MNEs such as the UK and Australian Diverted Profit Taxes and the recent US Base Erosion Anti-Abuse Tax.

France is the last country where a digital services tax (DST) has been introduced, despite the objections made by the United States. Starting from January 2019, the DST is imposed at a rate of 3% on the gross revenues derived from digital activities of which French "users" are deemed to play a major role in value creation. The law affects all digital business models which have registered during the previous calendar year revenue above one of the following two thresholds: €750 million for taxable digital services supplied worldwide; or €25 million for taxable digital services supplied in France.
Also Hungary has previously implemented a type of digital services tax, albeit only for advertising in Magyar language.
Austria, Belgium, the Czech Republic, Italy, Poland, Slovenia, Spain, and the United Kingdom have all either announced or published a proposal to introduce a DST[2].
Furthermore, worthy of note is the India's introduction of "significant economic presence" jointly with an equalization levy.
Of broader impact has been the United Kingdom's Diverted Profits Tax which aims at tackle profits that have been artificially diverted from the UK. Albeit it is not addressed to digital taxation, in fact, it can indirectly hit digital businesses[3].
Suchlike is the Australia's Multinational Anti-Avoidance Law (MAAL) and Australia's Diverted Profit Tax (DPT). The former is an anti-abuse rule limited in scope

[2]For a comprehensive review of actual state of rules see KPMG (2020)

[3]As regard to this, interesting is a study made by Liberini et al. (2020) which highlights the change in strategy by Facebook after the introduction of the UK's tax.

to non-resident enterprises belonging to large MNEs. Through penalty, akin to the cancellation of the tax benefit for MNE, the measure is targeted at deterring certain taxpayer behaviours, such as the use of trade structures involving remote sales of digital products and services. The latter is essentially designed to work as a deterrent and improve compliance with corporate tax rules and tax authorities.
Therefore, an increasing number of countries has unilaterally implemented a variety of measures aimed at protecting and/or expanding the tax base in the country where the customers or users are located.
On one side, all these initiatives are taken to increase the level of taxation of digitalized businesses, but on the other side they are likely to generate some economic distortions, double taxation problems, increased uncertainty and complexity, and associated compliance costs for businesses operating cross-border and, in some cases, they may potentially conflict with some existing bilateral tax treaties.
In conclusion, digitalization is a hot topic and it requires attention by different agents, among which economists who can contribute to the actual debate and shed light on pros and cons of different policies aimed to tackle digital challenges. In the next paragraph, we will account for the actual economic debate and the literature achievements reached so far.

1.2 Research area for economists

Digital economy constitutes a fruitful field for economists for several reasons.
First of all, defining a digital MNE as a company operates in multi-sided markets, where one group, for instance the users, affects the outcome for another group, such as the advertisers, across a platform through positive or negative externalities, it opens at the young branch of literature about two-sided platforms.
Their intrinsic characteristics like market power, economy of scale and economy of

scope, externalities, are all related to topics which are well-known to Industrial Organization.
Moreover, the role of users and their data have some implications on how to measure economic value and on privacy cost.
The importance of targeted ads as unique or main source of revenue for global big players revives the debate between informative versus persuasive advertising. In addition, personal data allow also an in-depth knowledge of final users, a better personal marketing campaign that ends in the capability for the sellers to discriminate the price not only among users but also at different time.
Behavioural economics looks at investigating social networks and their impact on different aspects of consumer's life through experiments, whose goal is to show the irrationality of choices and how these are leaded by emotions. The zero-price effect plays a key role, moving the consumer from an interior to a corner solution.
Furthermore, the absence of a link with a specific jurisdiction requires attention by international researchers. In particular, tax competition among jurisdictions, cross-borders issues, direct versus indirect taxation are all fields for empirical researches to analyse the efficacy of unilateral fiscal policies and their relative consequences on other countries.
Last but not least, the wealth concentration in a few big companies calls for distributional concerns and inequality issues. How to weight users' utility through a well-being function and the optimal tax problem are totally new research areas to be developed.

1.2.1 The current literature

The branch of industrial organization studying multi-sided markets is still young, and only in the last years it has been applied to digital businesses. It starts with the seminal contributions of Caillaud and Jullien (2003), Rochet and Tirole (2003 and

2006), and Armstrong (2006). Other contributions come recently by Belleflamme and Toulemonde (2009 and 2016), Belleflamme and Peiz (2010), Weyl (2010). All these models tend to emphasize mainly three features: the capability of a platform to provide distinct services to two sides of the market, charging different prices (multi-product firm); the (partial or total) internalization of the externalities present in the markets (both cross network effects and intra-group externalities); finally, the fact that platforms act as price setters (like monopolistic or oligopolistic, according to different markets and business models) on both sides of the market and typically (but not always) set uniform prices.

As for advertising issues, the dichotomy between informative and persuasive dates back to the last century. The economic analysis of advertising begins with Marshall (1890,1919) and Chamberlin (1933). Actually three views have emerged, with each view in turn being associated with distinct positive and normative application: advertising as informative, persuasive or complementary. We refer to Bagwell (2007) for a summary of the whole literature about advertising.

Here we look at the relevant role played by ads in catching consumers' attention and as main source of revenue for many digital businesses. Indeed, a growing number of platforms allows users to enjoy free services exposing them to ads in many ways and through several instruments, which in some cases blurs the edge between commercial and personal data. The so called "zero price effect" has been studied by Shampanier, Mazar and Ariely (2007), the importance of catching users' attention by Prat and Valletti (2019), the "privacy paradox" by Esayas (2018), according to which consumers express concern about their privacy but, in reality, they do very little to protect it and they are usually willing to disclose information in order to access to better quality services.

This last topic is related not only to economy, but in particular to law and psychology, therefore it requires a multidisciplinary approach.

Then we have a few papers that consider taxation in two-sided markets in general, and digital economy in particular, trying in this way to respond to the need of knowledge and research in how to tackle it, expressed also by supranational institutions, like the European Commission and the OECD. These works present different settings and different focus, but in some aspects, they are related to the issues already presented above.

Kind et al. (2010, 2009, 2008) are mainly concerned by comparing the impacts of ad valorem and unit taxes on fiscal revenues and on welfare in two-sided markets, with a specific focus on advertising-financed media. Kotsogiannis and Serfes (2010) adress the issue of taxation of two sided platforms in terms of tax competition between countries. Tremblay (2016) studies optimal taxation of a monopolist two-sided platform with two tax instruments, one on content and one on platform itself. Bloch and Demange (2017) focus on the effect of taxes on privacy protection, while Bourreau et al. (2017) assess the likely impacts of a tax on data collection and a tax on ads on the platform's business strategy and on fiscal revenues. Finally, Belleflamme and Toulemonde (2016) focus on the effects of taxation for competing two-sided platforms and how they are passed on pricing decisions.

All these works are theoretical and move into a monopoly or oligopoly framework. In fact, there are very few empirical papers due to the lack of available data and the fog around pricing strategy implemented by the biggest digital players.

1.2.2 Further future developments

There remains a very large number of research areas to investigate for economists, and not only, in the field of digitalization. There are many interesting elements that have not been investigated yet.

One aspect consists in the likely impact of unilateral choice made by a government on well-being in home and foreign country. It would constitute an additional chapter

to the broad literature on tax competition between countries, summed up by Keen and Konrad (2012).
The role played by data and the aftermaths on data mining on users constitute another chapter that has had too few attention.
How to build a well-being function which takes care of privacy, how preferences differ between individuals and social ones, what the sphere of action of a benevolent government is and how it can act, are all questions that need answers.

1.3 The present work

The present work constitutes a first attempt of understanding the digitalization of economy for what concerns platforms, which exploit two sides of market to make profits, by using targeted advertising built on personal "free" data.
In the following chapters we are going to show some different specifications to deal with two kinds of problem.
In the next Chapter, we analyse the nature of targeted advertising. We focus on the impact of ads on well-being of consumers. The possibility of having too much advertising calls for a government fiscal action that can be welfare-enhancing under some circumstances. Moreover, in a context of open economy, the decisions made by a government may spread their effects also abroad. We find out the condition under which this may happen.
In Chapter 3, we study in depth how a two-sided platform works. We introduce a new element to the previous chapter and to the related literature. Users care about their privacy but they do not know the actual level of their data mining. We deal with two issues: privacy protection and the need for government of raising fiscal revenue. First, we show how a change in privacy-consciousness is sufficient to affect platform's optimal choices. Second, we focus on the effects consequent to the introduction of some specific taxes.

Finally, in the last Chapter, we enlarge our approach by asking why a government should intervene by changing the market equilibrium. In an optimal taxation framework, we draw a method to address persuasive advertising when "undesired" by the government and show how it mixes with distributional concerns.
We are well-aware that our goal is ambitious and not easy to achieve, nevertheless our aim is rather to explore the "grammar of arguments" and to give some possible answers to the need of research into this area.

References

[1] Armstrong, M. (2006) *"Competition in Two-Sided Markets"*, RAND Journal of Economics, 37, 668-691.

[2] Bacache Beauvallet, M. (2017) *"Tax competition, tax coordination, and e-commerce"*, Journal of Public Economic Theory, 00, 1-18. https://doi.org/10.1111/jpet.12254

[3] Bagwell, K. (2007) *"The Economic Analysis of Advertising"*, Chapter 28 in Handbook of Industrial Organization, vol. 3, pp 1701-1844, Elsevier

[4] Belleflamme, P., and E. Toulemonde (2009) *"Negative intra-group externalities in two-sided markets"*, International Economic Review, 50, 245-272.

[5] Belleflamme, P., and E. Toulemonde (2016) *"Tax Incidence on Competing Two-Sided Platforms: Lucky Break or Double Jeopardy"*, CESifo Working Paper Series.

[6] Belleflamme, P., and M. Peiz (2010) *"Platform competition and seller investment incentives"*, European Economic Review, 54, 1059-1076.

[7] Bloch, F., and G. Demange (2017) *"Taxation and privacy protection on Internet platforms"*, Journal of Public Economic Theory, 00:1-15. https://doi.org/10.1111/jpet.12243

[8] Bourreau M., B. Caillaud, and R. DeNijs (2017) *"Taxation of a digital monopoly platform"*, Journal of Public Economic Theory, 00:1-12. https://doi.org/10.1111/jpet.12255

[9] Chamberlin, E. (1933) *"The Theory of Monopolistic Competition"*, Cambridge, MA: Harvard University Press.

[10] Caillaud, B., and B. Jullien (2003) *"Chicken and egg: Competition among intermediation service providers"*, RAND Journal of Economics, 34, 309-328.

[11] Colin, N., and P. Collin (2013) *"Mission d'expertise sur la fiscalité de l'économie numérique"*, Paris. Retrieved from:
http://www.economie.gouv.fr/files/rapport-fiscalite-du-numerique_2013.pdf

[12] Esayas, S. Y. (2018) *"Competition in (data) privacy:'zero'-price markets, market power, and the role of competition law"*, International Data Privacy Law, 8(3), 181-199.

[13] European Commission (2018) *"Digital Taxation: Commission proposes new measures to ensure that all companies pay fair tax in the EU"*, Brussels, 21 March 2018.

[14] Gaspar, V., P. Collin, M. P. Devereux, J. H. Snabe, T. Varrak, M. Walsh, and B. Westberg (2014) *"Expert group on taxation of the digital economy"*, Brussels: European Commission. Retrived from:
http://ec.europa.eu/taxation_customs/business/company-tax/tax-good-governance/expert-group-taxation-digital-economy_en

[15] Keen, M., and K. A. Konrad (2012) *"The Theory of International Tax Competition and Coordination"*, Max Planck Institute for Tax Law and Public Finance.

[16] Kind, H.J., M. Koethenbuerger, and G. Schjelderup (2008) *"Efficiency enhancing taxation in two-sided markets"*, Journal of Public Economics, 92, 1531-1539.

[17] Kind, H.J., M. Koethenbuerger, and G. Schjelderup (2009) *"On revenue and welfare dominance of ad valorem taxes in two-sided markets"*, Economics Letters, 104(2), 86-88.

[18] Kind, H.J., M. Koethenbuerger, and G. Schjelderup (2010) *"Tax responses in platform industries"*, Oxford Economic Papers, 62(4), 764-783.

[19] Kotsogiannis, C., and K. Serfes (2010) *"Public goods and tax competition in a two-sided market"*, Journal of Public Economic Theory, 12.2: 281-321.

[20] KPMG (2020) *"Taxation of the digitalized economy"*, April 4th, 2020. Retrieved from:
https://tax.kpmg.us/content/dam/tax/en/pdfs/2020/digitalized-economy-taxation-developments-summary.pdf

[21] Liberini, F., Á. Cuevas, R. Cuevas, A. Lassmann, and A. Russo (2020) *"Taxation and Global Spillovers in the Digital Advertising Market. Theory and Evidence from Facebook."*, Working paper.

[22] Marshall, A. (1890) *"Principles of Economics"*, London: MacMillan and Co..

[23] Marshall, A. (1919) *"Industry and Trade: A Study of Industrial Technique and Business Organization; and of Their Influences on the Conditions of Various Classes and Nations"*, London: MacMillan and Co.

[24] OECD (2014) *"Report on the tax challenges of the digital economy"*, Paris: Center for Tax Policy and Administration, Committee on Fiscal Affairs.

[25] OECD (2018) *"Tax Challenges Arising from digitalization - Interim Report 2018: Inclusive Framework on BEPS"*, OECD/G20 Base Erosion and Profit Shifting Project, OECD Publishing, Paris.

[26] OECD (2019) *"Addressing the Tax Challenges of the digitalization of the Economy"*, Public Consultation Document OECD/G20 Base Erosion and Profit Shifting Project, OECD Publishing, Paris, 13 February - 6 March.

[27] Prat, A., and T. M. Valletti (2019) *"Attention Oligopoly"*, SSRN Electronic Journal.

[28] Rochet, J.C., and J. Tirole (2003) *"Platform competition in two-sided markets"*, Journal of the European Economic Association, 1, 990-1029.

[29] Rochet, J.C., and J. Tirole (2006) *"Two-Sided Markets: A Progress Report"*, Rand Journal of Economics, 6, 417-429.

[30] Russo, J.C., and J. Tirole (2006) *"Two-Sided Markets: A Progress Report"*, Rand Journal of Economics, 6, 417-429.

[31] Shampanier, K., N. Mazar, and D. Ariely (2007) *"Zero as a Special Price: The True Value of Free Products"*, Marketing Science, 26(6), 742–757.

[32] The Economist (2018) *"Why is Mark Zuckerberg testifying in Congress?"*, April 9th, 2018. Retrieved from: https://www.economist.com/the-economist-explains/2018/04/09/why-is-mark-zuckerberg-testifying-in-congress

[33] Tremblay, M. J. (2016) *"Taxation on a Two-Sided Platform"*, Mimeo.

[34] Weyl, E.G. (2010) *"A Price Theory of Multi-Sided Platforms"*, American Economic Review 100, 1642-1672.

Chapter 2

Targeted Advertising on the web: too little or too much?

This paper addresses the question of the targeted web advertising and its impact on consumers. By assuming the existence of a monopolistic digital platform that can link two sides of the market, we find out the condition under which production choices are far from the optimal ones. The possibility of having too much advertising calls for a government fiscal action that can be welfare-enhancing under some circumstances. Moreover, in a context of open economy the decisions made by a country could spread their effects also abroad.

JEL classification: D42, H22, M37.

Keywords: digital platform, turnover tax, unit taxes, targeted advertising, monopoly.

2.1 Introduction

The digitalization has allowed to connect different sides of the market through the development of new businesses.

Almost everywhere two-sided firms serve distinct customer groups that are connected through interdependent demand. A classic example is a platform, whose revenues accrue from advertising displayed on a web site and from access fee paid by people to use services on that site, such as videos or streaming platforms, social networks, etc.

These businesses are characterized by quantity spillovers as platforms maximise profits by facilitating value-creating interactions between groups of agents.

Moreover, in the last years, the standard two-sided business model has implemented a novelty regarding the role of data. In fact, most of the value created by a digital platform comes from the data exploitation, though this input is provided by users for free.

There exist several policy issues related to the taxation of firms operating within the digital sector. It is possible to detect two opposing trends: one aims at maximizing collections based on exponentially growing digital flows; the second one recognizes instead that lowering taxation benefits consumers and businesses, and consequently, economic growth. The two-sided or multi-sided nature of Internet platforms engaged in digital advertising presents a greater number of taxation challenges since the pricing model of digital advertisers involves at least three parties: the publisher, the advertiser, and the user; each of which could be located in different jurisdictions or countries. For this reason, an emerging view posits that online advertising should be taxed in three jurisdictions.

The present analysis is in line with recent papers in Industrial Organization which study two-sidedness (Anderson and Coate, 2005; and Rochet and Tirole, 2003, 2006) and how taxation may impact on markets (Kind et al., 2008, 2009, 2010). It also looks at the data question investigated by Bourreau et al. (2016) and Bloch and Demange (2016).

No one of these has attempted to abandon the concept of a closed economy to see what might be the effects of taxes in a context with more countries.

For what we know there exist only two working papers, one by Cui and Hashimizade (2019), which studies the likely outcome of a digital service tax with a theoretical model, and another one by Liberini and al. (2020), which is the first attempt of making an empirical analysis of Facebook ads' prices.

Furthermore, also advertising literature plays a crucial role (Dixit and Norman, 1978; Shaphiro, 1980). Here, we look at advertising as a tool for catching consumers' attention and as main source of revenue for many digital businesses (such as Google or Facebook).

The present work mainly seeks to answer the following question: in a context where a country needs to increase its fiscal revenue, how platform's taxation may impact the welfare for all involved agents, and whether there might be effects abroad.

By using a theoretical model, we keep in mind the following main features while looking for an answer: a platform behaves like a monopolist, there exist difficulties in imposing direct taxes on platform's revenue, a crucial role is played by data exploitation in value chain, and users' information is bought for free, so we can talk of “free work".

The rest of the paper proceeds as follows: in Section 2, we present the model; then, in Section 3, we introduce the issues related to the taxation of digital economy; in Section 4, we look at the aftermaths in a foreign country caused by some domestic measures; and, finally, we conclude with some comments and possible future insights.

2.2 The model

Consider a digital platform placed abroad which allows a domestic firm to sell a good x to a population of consumers in the same country. One can think of an e-commerce shop as the retailer, but also traditional businesses buy ads' space and sell products online.

The role of the platform is not only that of linking the two sides of the market, it also sells a targeted advertising service to the retailer. Hence, it builds a two-sided market by exchanging on one side ads directly and on the other side the good x indirectly.

The reason why we set the platform abroad will be clear later.

Here we are not considering how many users will be addressed by the platform and why they choose to be online, so their mass is fixed and normalized to one.

The digital platform builds the advertising package as a combination of two factors: ads "intensity", a, and quality, captured by a proxy, s, for the level of data exploitation measured by the exchanged bits for instance. So, thereafter we use the term "data exploitation" and "quality" with the same meaning. This package is sold to a domestic firm at the price A through a system of auctions which allows the platform to take the whole producer's profit.

There exist lots of examples which endorse this scheme. For instance, let us think about Amazon based in US that connects retailers in UK to customers in India. Or a Google branch established in Ireland that makes profits all over Europe, and allows retailers to buy ads' spaces and "sponsoring" associated with research keys through a complex system of auctions.

In our model, the domestic retailer sells good x at price p and can exploit its monopolistic power thanks to the auction since this allows it to be the only seller on the market. Implicitly we are assuming that there is another potential entrant on the market but only one of them can produce and sell to consumers. However, the firm cannot discriminate among customers due to arbitrage opportunities, i.e. the good x can be resold between them. The demand for the product is $D(p, a, s)$, and the inverse demand is $P(x, a, s)$. Costs of producing x units are $c(x)$, and they are increasing in the produced quantity at a non decreasing rate, that is $c_x > 0$ and $c_{xx} \geq 0$, where the subscript indicates the derivative. Hence, retailer's profits are:

$$\Pi^x = px - c(x) - A \tag{2.1}$$

while the platform's profits are:

$$\Pi^p = A - k(a, s) \tag{2.2}$$

where the function cost $k(a, s)$ is non decreasing and convex in its arguments, that is $k_a \geq 0$, $k_{aa} \geq 0$, $k_s \geq 0$, and $k_{ss} \geq 0$.

The use of a unique price for the package of tailored ads allows the platform to avoid the problem of double-marginalization and to appropriate the retailer's profit by imposing a "fee" equal to the vertical structure's profit.:

$$A = \overline{\Pi}^x(p^m, a, s) - \Pi^x(p^*, 0, 0) \tag{2.3}$$

where p^m is the monopolistic price, and p^* is the competitive price[1]. Here, we are excluding any risk-aversion by the retailer or uncertainty about the final demand.

Therefore, the game structure is similar to Stackelberg's one: the platform anticipates the firm's action by choosing the level of ads a and data exploitation s that maximises the retailer's profit. Then, the firm sets the monopolistic price $p^m(a, s)$ and pays the "fee" A, earning zero. Customers buy the good x according to the price p^m.

[1]Note that the profit $\overline{\Pi}^x$ is without "fee", and it corresponds to the maximum earnings for the retailer given the level of a and s chosen by the platform. The profit without advertising is zero if the retailer plays in a perfectly competitive market and the marginal cost is constant.

Our goal is to understand the conditions under which the optimal choice of targeted advertising made by the platform is not socially efficient.

The retailer has an incentive to raise the targeted advertising if this increases the marginal willingness to pay of the marginal consumer, that is $x\Delta p(x)/\Delta i > \Delta k/\Delta i$ with $i = a, s$. But this is not an accurate measure of the social benefits coming from the rise since we have to consider not the marginal consumer but the average marginal effect on the entire population. The targeted ads increase is desirable only if the average benefit exceeds the average cost, i.e. if $(1/x)\int_0^x \Delta p(v)dv > \Delta k/x$. This result is quite similar to that found by Spence (1975) about under/over-provision of quality in monopoly.

The average benefit coincides with the increase in revenue for the retailer only when the marginal consumer is representative of the entire population, that is to say when $(1/x)\int_0^x \Delta p(v)dv = \Delta p(x)$. In what follows, we are going to show when this equality does not hold.

First, let us focus on the Consumer surplus which can be expressed in two ways:

$$S = \int_0^x P(v, a, s)dv - xP(x, a, s) \tag{2.4}$$

or, equivalently as:

$$S = \int_p^\infty D(v, a, s)dv \tag{2.5}$$

Profits for the retailer are always equal to zero, whereas profits for the platform can be re-expressed by using (2.1) and (2.2) as:

$$\Pi^p = p(x, a, s)x - c(x) - k(a, s) \tag{2.6}$$

whose first order conditions (FOCs) are:

$$p - c_x = -xp_x \tag{2.7}$$

$$xp_a - k_a = 0 \tag{2.8}$$

$$xp_s - k_s = 0 \tag{2.9}$$

The reader should notice that the condition (2.7) is set only after the optimal choice of a and s, and it is not chosen by the platform jointly with the others. Due to this reason, the equilibrium of a fully integrated platform might differ.

Looking carefully at conditions (2.8) and (2.9), it is quite evident that the ratio between level of ads and data exploitation is always efficiently set by the digital platform whatever the provided quantity x, that is:

$$\frac{p_a}{p_s} = \frac{k_a}{k_s} \tag{2.10}$$

The platform chooses directly two of four variables: quantity of ads and data exploitation; the price is chosen by the maximizing retailer, whereas the fourth, the quantity of good x is determined by the demand function.

Making a comparison between consumer surplus and profit, it is apparent that these two may differ. For a given quantity, x, a social planner who wants to maximize the consumer surplus with respect to ads, a, and data, s, has to satisfy these two conditions:

$$\frac{\partial S}{\partial a} = \int_0^x p_a dv - k_a = 0 \tag{2.11}$$

$$\frac{\partial S}{\partial s} = \int_0^x p_s dv - k_s = 0 \tag{2.12}$$

Hence the level of ads and data differs from the socially efficient ones for the consumers each time that the integrals are not equal to the marginal valuation of targeted advertising. In other words, by comparing expressions (2.11) and (2.12) with

(2.8) and (2.9), an increase in consumer surplus depends on the relative magnitudes of $\int_0^x p_a dv$ to xp_a and of $\int_0^x p_s dv$ to xp_s. By dividing both terms by x, we obtain the average valuation respectively of ads and data at the margin over the whole population in the market. So, when the increase in this last exceeds the marginal valuation, the level of ads and data is too low compared to the optimum, otherwise it is set too high.

Proposition 1 *For given x, the platform oversupplies advertising relative to the optimum when $\frac{1}{x}\int_0^x p_a dv < p_a$ and conversely. A sufficient condition for over-provision is that the marginal value of ads increases as absolute willingness to pay falls, that is $p_{xa} > 0$, and conversely for under-provision.*

A caveat of the proposition above is that it tells us which way the platform biases its choice of ads only if the output is the same both for the monopolist and for the social planner.

Proposition 2 *For given x, the platform overexploits data relative to the optimum level when $\frac{1}{x}\int_0^x p_s dv < p_s$ and conversely. A sufficient condition for over-exploitation is that the marginal value of data increases as absolute willingness to pay falls, that is $p_{xs} > 0$, and conversely for under-exploitation.*

The caveat is always the same. When ads quantity and quality are complementary, that means that the marginal value of data increases in ads, i.e. $p_{as} > 0$, for a given x, the platform ends up oversupplying targeted advertising. An increase in one variable strengthens the effect of the other at the margin.
In other words, the term p_{xa} means that the effect of an increase in ads is greater at the margin when the willingness to pay declines, while p_{xs} stands for a greater effect of an increase in data exploitation at the margin when the willingness to pay declines. Therefore, since the platform, being a monopoly, cares only about what happens at the margin, for a little shift of monopoly's quantity with respect to social

optimum, it will produce too much targeted advertising.

Then, one should notice that the condition (2.10) is quite different from the social optimum ratio expressed as follows:

$$\frac{\int_0^x p_a dv}{\int_0^x p_s dv} = \frac{k_a}{k_s} \tag{2.13}$$

Since $x^* > x^M$ at the two solutions, nothing guarantees that the left sides of (2.10) and (2.13) coincide. Though efficient, platform's choices about how much to invest on ads or data will end by not being the social preferable ones.

Thus, the digital firm with market power deviates in three aspects. First, it sets a markup above the marginal cost, selling too little quantity of good x. Second, the amount of ads may be too high or too low according to the relative valuations of ads of the marginal and average consumer. Third, the level of data exploitation may not be optimal depending on the relative valuations of "privacy" of the marginal and average customer.

Another interesting circumstance to investigate is when the retailer is committed to set the price at the competitive level, that is $p(x, a, s) = c_x$.

Under this circumstance, a distortion would still happen. Indeed, the social planner aims at maximizing the area between the demand curve and the marginal cost of providing the quantity, whereas the digital platform focuses on maximizing the area between the final price and the marginal cost. When this latter area grows faster than the former one as targeted advertising goes up, then the monopolist oversupplies ads.

On the contrary, when the marginal cost increases slowly, in a way that is quasi-horizontal, it is possible that the consumers' surplus grows faster than the platform's profit. In this case, the firm undersupplies targeted ads[2]. The present result is quite

[2]To see it, let us assume that $c_x = 0$, such as the cost is constant. It is quite simple to show that when the maximum willingness to pay does not change, that is the intercept on the vertical axis is fixed while the slope of the demand changes in relation to a and s, consumers' surplus increases with them. So any

similar to that one obtained by Kind et al. (2008).

Coming back to monopoly case, generally the relationship between optimal and profit maximising targeted advertising is determined by (i) the interaction of the average marginal effect and (ii) the extent to which the monopolist restricts output. This second factor can be caught also by the shape of the demand curve, or roughly by the elasticity of demand.

Intuitively, one should investigate how the fraction of total potential surplus catchable by the firm varies with targeted ads. It depends on any demand function $D(p,a,s)$, except when the elasticity is independent of price.

Proposition 3 *If the price elasticity is not a function of the price and if the elasticity increases with targeted advertising, then ads are oversupplied and conversely. If the elasticity does not vary with targeted advertising, then ads level and data exploitation are set at the optimal level by the monopolist.*

To prove it[3], let us suppose that the demand function is a CES type, for any given level of targeted advertising, so that $P(x,a,s) = g(a,s)x^{-n(a,s)}$, where $g(a,s)$ is a generic function and $n(a,s)$ is the inverse of the price elasticity of demand.

Moreover, consider the marginal cost of producing an additional unit constant and independent of the level of a and s. Define, in addition, the term β as the ratio of maximized profits to maximized surplus, that is:

$$\beta(a,s) = \frac{\overline{\pi}(a,s)}{\overline{W}(a,s)}$$

contraction in a and s, by decreasing the elasticity of demand, shrinks the area of consumers' surplus triangle.

[3]The demonstration is similar to that one used by Spence (1975).

It is possible to show[4] that:

$$\beta(a,s) = [1 - n(a,s)]^{\frac{1}{n(a,s)}} \tag{2.14}$$

and, as consequence, if price elasticity increases with targeted advertising, that is with a and s, then $n'(a,s) < 0$ and $d\beta/dn < 0$[5]. Since in the optimum for the monopolist, we have $\pi' = 0$, then $W' < 0$, which means that targeted ads are too many from a social planner prospective.

Hence, the above mentioned failures of market open to government's intervention in order to correct them. In next paragraph, we show how some kinds of taxes may impact the equilibrium.

2.3 Taxation

In this section, we want to show how different implemented fiscal instruments might generate divergent effects.

We suppose to use four kinds of taxes: a tax on turnover, labelled as t_r, an ad valorem tax, τ, a specific tax on ads or on data[6], t_a and t_s respectively.

---

[4]Starting from the profit function for the monopolist: $\pi = P(x,a,s)x - c(a,s)x$; the quantity chosen is equal to $x^m = \left[\frac{c}{g(1-n)}\right]^{-\frac{1}{n}}$. By substituting it into the monopolist's profit function, we get: $\pi = g^{\frac{1}{n}} c^{-\frac{1-n}{n}} (1-n)^{\frac{1}{n}} \left[(1-n)^{-1} - 1\right]$.
Doing the same for the social planner function, that is: $W = \int_0^x P(v,a,s) - c(a,s)x$, we find the optimum quantity $x^* = \left(\frac{c}{g}\right)^{-\frac{1}{n}}$. So, the welfare function becomes: $W = g^{\frac{1}{n}} c^{-\frac{1-n}{n}} \left[(1-n)^{-1} - 1\right]$.
As a consequence, it follows that the ratio π/W is equal to: $\beta = (1-n)^{\frac{1}{n}}$.

[5]Taking the derivative of β with respect to n, after some calculations, we get:

$$\frac{d\beta}{dn} = \left[-\frac{1}{(1-n)n} - \frac{ln(1-n)}{n^2}\right](1-n)^{\frac{1}{n}}$$

which is negative as long as $n < 1$.

[6]In a French study on the taxation of the digital economy by Colin and Collin (2013), the authors talk about implementing a tax based on the effectively free use of personal data provided by internet users which businesses use as means of creating value, in addition to three proposed new taxes, including a tax on on-line advertising, a tax on e-commerce services and the extension of the current tax on the sale of rental of videos and films including video on demand to foreign internet players.

Looking at the profit for the platform, the introduction of a tax changes the expression (2.2) and the relative first order conditions. When we use a tax on revenue, we get:

$$\Pi^{p_r} = (1 - t_r)A - k(a, s) \tag{2.15}$$

where the subscript indicates the fiscal instruments. It is possible to show that the efficiency ratio does not change neither under a tax on revenue nor with a VAT.
Instead, using a specific tax either on ads or on data, the condition (2.10) becomes:

$$\frac{p_a}{p_s} = \frac{k_a + t_a}{k_s + t_s} \tag{2.16}$$

Using comparative statics, under whatever taxation, the amount of ads and data decreases jointly[7].
Using a utilitarian welfare function, when the platform is located abroad, the government takes into account only the fiscal effects on consumers' surplus and on firm's profits, which are zero under our specifications:

$$W = S + \Pi^x + T = \int_0^x p(v, a, s)dv - p(x, a, s)x + T \tag{2.17}$$

where the term T indicates the fiscal revenues accruing from taxation.
Now, according to a general instrument t, the welfare variations result as:

$$\frac{dW}{dt} = -x\frac{dp}{dt} + \frac{da}{dt}\int_0^x p_a dv + \frac{ds}{dt}\int_0^x p_s dv + \frac{dT}{dt} \tag{2.18}$$

Being the fiscal revenue effect always positive when evaluated in $t = 0$, the above condition calls undoubtedly for an increase in welfare when the price decreases and this reduction times the number of unit consumed is greater than the loss for marginal consumers due to the reduction in targeted advertising. Looking at the effect of taxes on ads and data exploitation, we know that they are negative from

[7] The whole derivation is shown in the Appendix.

comparative statics.

Graphically, Figure 2.1 shows well why the difference in size of the two areas determines if the tax is welfare enhancing or not. Indeed, when the area in blue is smaller than the area in purple, then the welfare increases for sure.

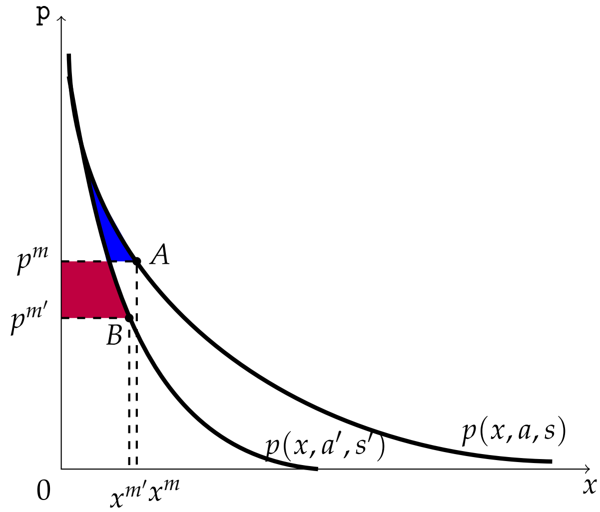

FIGURE 2.1: Welfare's comparison

The essential assumption is that the strength of the effect of targeted advertising is decreasing in consumers' marginal willingness to pay. It corresponds to assume that marginal willingness to pay rises with targeted ads, i.e. $p_{xa} > 0$ and $p_{xs} > 0$. As a consequence, after the introduction of taxes the platform reduces both ads and data exploitation from a and s to a' and s' respectively, and the demand curve shrinks inward. The equilibrium shifts from point A to B, where both price and quantity are lower. Implicitly, we are assuming that the platform overproduced targeted advertising, and despite the distortion induced by the monopolist's power of the retailer continues to hold, this distortion reduces when advertising falls.

Next step consists of studying which of the previous kinds of taxes may be more efficient. After having implemented a tax, a government causes a distortion to the equilibrium in exchange of a positive fiscal gain.

Any indirect tax generates some distortion, but there are some whose effect is greater in terms of distortion or lower in terms of fiscal revenue. We seek to identify in this way the tax that generates the highest fiscal gain, with the same distortion.

In order to do it, we start by comparing the first order conditions in presence of taxes[8]. When these conditions are identical under two different taxes, then the quantity choices are the same and so it is the distortion. By comparing the fiscal revenue it is hence possible to find the best fiscal instrument.

Let us begin by comparing two specific taxes on advertising and data and a tax on revenues:

$$(1-t_r)xp_a = xp_a - t_a \rightarrow t_a = t_r x p_a \tag{2.19}$$

$$(1-t_r)xp_s = xp_s - t_s \rightarrow t_s = t_r x p_s \tag{2.20}$$

Using the above expressions, let us define tax revenue under respectively turnover and unit taxation as:

$$R^{rev} = t_r A = t_r[px - c(x)] \tag{2.21}$$

$$R^{unit} = t_a a + t_s s = t_r x(p_a + p_s) \tag{2.22}$$

Taking the difference between expressions (2.21) and (2.22), we obtain:

$$R^{rev} - R^{unit} = t_r[px - c(x) - (p_a + p_s)x] \tag{2.23}$$

whose sign depends on the strength of marginal effect of a and s onto price p.

Proposition 4 *A turnover tax generates more fiscal revenues compared to specific taxes when the price minus the joint marginal effect of targeted advertising of a and s on p is greater than the average cost, that is $(p - p_a - p_s) > \frac{c(x)}{x}$, and conversely.*

[8] The reader can find easily the first order conditions for each kind of taxes in the appendices.

Making the same exercise by comparing a VAT and two specific taxes, we obtain:

$$R^{vat} = \frac{\tau}{1+\tau}A = \frac{\tau}{1+\tau}[px - c(x)] \quad (2.24)$$

$$R^{unit} = t_a a + t_s s = \frac{\tau}{1+\tau}x(p_a + p_s) \quad (2.25)$$

Taking again the difference between expressions (2.24) and (2.25), we obtain:

$$R^{vat} - R^{unit} = \frac{\tau}{1+\tau}[px - c(x) - (p_a + p_s)x] \quad (2.26)$$

whose sign always depends on the strength of marginal effect of a and s onto price p. When the square bracket is negative, then there exists a pair of specific taxes which revenue-dominates an ad valorem tax.

Proposition 5 *An ad valorem tax generates more revenue than specific taxes when the per-unit advertising expenditure is greater than the sum of marginal effects of targeted advertising, that is* $\frac{A}{x} = p - \frac{c(x)}{x} > p_a + p_s$*, and conversely.*

The interpretation of this result tells us that if we start from an equilibrium with a positive tax, for instance $\tau > 0$, and the marginal effects of p_a and p_s are very high, then we can do better by substituting the VAT with two specific taxes. In fact, by substituting the ad valorem tax with two specific taxes, the government has to impose a higher unit tax rate to keep output unchanged or otherwise it may reduce the fiscal burden and consequently the tax distortion by levying the same tax revenue.

2.4 The model for an open economy

Now, we will seek to point out the feasible indirect consequences that a unilateral action played by a domestic government may have on a foreign country's economy. We want to figure out how the effects of taxes could be spread from the home country to the foreign one.

We consider here the existence of two countries, a domestic one labelled by h and a foreign one, labelled by f. The platform is located out of these two countries so it cannot be affected by standard corporate tax rate, but only through indirect taxation. The retailer is placed in home country h, but it is able to sell goods at home and abroad through the digital platform without having a physical presence in the foreign market. It is what happens actually where we think about digital services for instance.

The platform sets a distinct fee according to the destination market:

$$\Pi^p = A^h + A^f - k(a,s) \tag{2.27}$$

where total costs depend on total ads "intensity" and "quality", that is $a = a^h + a^f$ and $s = s^h + s^f$.

Consequently the retailer's profits function changes as follows:

$$\Pi^x = p^h(x^h, a^h, s^h)x^h + p^f(x^f, a^f, s^f)x^f - c(x) - A^h - A^f \tag{2.28}$$

where prices can be different between the two markets, and consumers cannot resell goods abroad, so we are assuming that there is no cross-border trade.

Looking at the first order conditions for the retailer, we get:

$$p^h - c_x \frac{\partial x}{\partial x^h} = -x^h p_x^h \tag{2.29}$$

$$p^f - c_x \frac{\partial x}{\partial x^f} = -x^f p_x^f \tag{2.30}$$

which is the standard distortion in a monopoly market, similar to that found in the closed economy.

Given the retailer's chosen prices, using expression (2.27), the platform maximises its profit function by choosing a and s for both markets:

$$\Pi^p = p^h(x^h, a^h, s^h)x^h + p^f(x^f, a^f, s^f)x^f - c(x) - k(a,s) \tag{2.31}$$

whose Focs are, taking into account that $\partial a/\partial a^i = \partial s/\partial s^i = 1$ with $i = h, f$:

$$x^h p_a^h - k_a = 0 \quad (2.32)$$

$$x^h p_s^h - k_s = 0 \quad (2.33)$$

$$x^h p_a^f - k_a = 0 \quad (2.34)$$

$$x^h p_s^f - k_s = 0 \quad (2.35)$$

And, by combining these expressions, it is possible to show that in the optimum the following condition must hold:

$$\left(\frac{p_a}{p_s}\right)^h = \left(\frac{p_a}{p_s}\right)^f \quad (2.36)$$

It tells us that the ratio between “intensity" and “quality" has to be the same at home and abroad. After the unilateral choice of the domestic government about setting an indirect tax, this ratio however may change and, consequently, effects may arise not only in the home country, but also in the foreign one.

2.4.1 Effects of home taxes

First, we look at a tax on turnover. Platform’s profit function changes as follows:

$$\Pi^{p_r} = (1 - t_r)A^h + A^f - k(a, s)$$

which can be rewritten as:

$$\Pi^{p_r} = (1 - t_r)\left[p^h x^h - c(x)\frac{x^h}{x}\right] + p^f x^f - c(x)\frac{x^f}{x} - k(a, s) \quad (2.37)$$

whose Focs are:

$$
\begin{aligned}
(1-t_r)x^h p_a^h - k_a &= 0 \\
(1-t_r)x^h p_s^h - k_s &= 0 \\
x^f p_a^f - k_a &= 0 \\
x^f p_s^f - k_s &= 0
\end{aligned}
$$

What happens is strictly clear. The level of a and s decreases in home country but through the cost function this effect is spread also abroad.

To derive a not ambiguous sign due to the complexity of using a matrix 4×4, let us assume thereafter that the marginal cost of ads "intensity" is constant and normalized to zero, i.e. $k_a = 0$.

Under this assumption and using comparative statics[9], one can show that:

$$
\begin{aligned}
\frac{da^f}{dt_r} &< 0 \\
\frac{ds^f}{dt_r} &< 0
\end{aligned}
$$

The same happens also when we use a specific tax on ads, t_a, or on data, t_s. Platform's profit function changes as follows:

$$\Pi^{p_r} = p^h x^h + p^f x^f - c(x) - k(a,s) - t_a a^h - t_s s^h \tag{2.38}$$

whose comparative static analysis gives the following results:

$$
\begin{aligned}
\frac{da^f}{dt_a} < 0 \quad &\text{and} \quad \frac{ds^f}{dt_a} < 0 \\
\frac{da^f}{dt_s} < 0 \quad &\text{and} \quad \frac{ds^f}{dt_s} < 0
\end{aligned}
$$

[9]See the Appendix A.5 for the whole derivation.

So the effect is always a contraction on the provision of ads and on data exploitation[10].

Therefore, the home negative effect of taxes on the two variables chosen by the platform can be mitigated and spread through the costs' channel also abroad. The reader should also notice that what we said before about the welfare-enhancing power of taxes still holds both at home and abroad. Though, in a foreign country, the final effect might be lower due to the lack of fiscal revenue effect which instead is present in the domestic country.

In fact, looking at the foreign well-being, we have:

$$\frac{dW^f}{dt_i} = -x^f \frac{dp^f}{dt_i} + \frac{da^f}{dt_i} \int_0^{x^f} p_a^f dv + \frac{ds^f}{dt_i} \int_0^{x^f} p_s^f dv \tag{2.39}$$

where the subscript $i = r, a, s$ indicates the kind of implemented tax.

The final result depends on the change of price for final users, when they pay more, the sign is unambiguous and the foreign welfare decreases, otherwise when it is negative, then the sign of final effect abroad depends on the strength of the "price effect" with respect to the targeted "advertising effect".

Hence this constitutes the only circumstance under which it is possible an improvement in the foreign well-being. However this gain is more difficult to be achieved with respect to the domestic improvement.

2.5 Summary and concluding remarks

The digitalization of the economy represents a big challenge for many actors in traditional and new markets, and also for fiscal authorities. The role of government has been studied looking at traditional forms of taxation and the feasible effects onto final consumers. One omitted aspect in the literature consists of the possibility that unilateral choice of a government might affect welfare besides home borders

[10]The whole derivation is shown in Appendix A.6.

through international links built by a digital platform.
The present study seeks to show how this situation can verify and to forecast likely effects resulting after the introduction of some kinds of taxes.
Through the common cost structure a two-sided platform is able to shift a slice of its fiscal burden not only onto the other side of the market, as shown in literature, but also on markets abroad, in a context of more countries.
Another aspect that we have learnt is that each fiscal action needs to evaluate first the impact of targeted advertising on final consumers' well being and then whether ads and data are under/over provided with respect to social optimum levels.
A fiscal action implemented by a country may generate a welfare improvement, when the platform is located abroad and the attention is focused only on domestic consumers. Nevertheless, the final effects overcome home borders and might impact also foreign consumers' well-being. A gain in welfare is not obvious and needs specific conditions.
For this reason, the choice of European Commission and OECD to follow a common path to face the problem, avoiding to let each state to act alone, seems to be the best way. Though it is not enough.
We showed also that the means chosen to rise fiscal revenues are relevant as well, and there are some of them that can be more effective reducing distortions.
This said, we know very well the drawbacks and the limits of our work, but at the same time, it can be considered a good starting point with lots of hints to examine in depth the issues and challenges that digitalization has opened.
As argued by Fuest (2018), looking at the whole picture, there are two additional issues to consider talking about taxing digital economy and shifting from origin to destination principle. Firstly, lower tax burden for digital businesses derives partially from tax credit for R&D activities. All relative benefits would be harmed by any taxes: "removing the preferential treatment that arises through tax subsidies for

research would be economically damaging as a result". The second point to stress concerns the proposal to allocate more taxation rights to those countries in which products are sold. As regard to, there is little empiric evidence to suggest that Europe would benefit from such a reform, in fact it is a net exporter.

Therefore, forthcoming papers will focus on other theoretical aspects neglected so far to embody the dynamism of investment in R&D for instance or different market competition structure for domestic retailers. However, empirical evaluations of unilateral tax reforms already set by single countries would help more to understand what happens behind the fog of multinational digital companies.

References

[1] Anderson, P.S., and S. Coate (2005) *"Market provision of broadcasting: a welfare analysis"*, Review of Economic Studies, 72, 947-972.

[2] Armstrong, M. (2006) *"Competition in Two-Sided Markets"*, RAND Journal of Economics, 37, 668-691.

[3] Bacache Beauvallet, M. (2017) *"Tax competition, tax coordination, and e-commerce"*, Journal of Public Economic Theory, 00, 1-18. https://doi.org/10.1111/jpet.12254

[4] Bagwell, K. (2007) *"The Economic Analysis of Advertising"*, Chapter 28 in Handbook of Industrial Organization, vol. 3, pp 1701-1844, Elsevier

[5] Bloch, F., and G. Demange (2017) *"Taxation and privacy protection on Internet platforms"*, Journal of Public Economic Theory, 00:1-15. https://doi.org/10.1111/jpet.12243

[6] Bourreau, M., B. Caillaud, and R. DeNijs (2017) *"Taxation of a digital monopoly platform"*, Journal of Public Economic Theory, 00:1-12. https://doi.org/10.1111/jpet.12255

[7] Bourreau, M., B. Caillaud, and R. DeNijs (2016) *"Taxation of a Digital Monopoly Platform"*, HAL archives-ouvertes, October 25.

[8] Chamberlin, E. (1933) *"The Theory of Monopolistic Competition"*, Cambridge, MA: Harvard University Press.

[9] Colin, N., and P. Collin (2013) *"Mission d'expertise sur la fiscalité de l'économie numérique"*, Paris. Retrieved from: http://www.economie.gouv.fr/files/rapport-fiscalite-du-numerique_2013.pdf

[10] Dorfman, R., and P.O. Steiner (1954) *"Optimal Advertising and Optimal Quality"*, America Economic Review, 44, 826-36.

[11] Cui, W., and N. Hashimzade (2019) *"The Digital Service Tax as a Tax on Location-Specific Rent"*, CESifo Working Paper Series No. 7737.

[12] Esayas, S. Y. (2018) *"Competition in (data) privacy:'zero'-price markets, market power, and the role of competition law"*, International Data Privacy Law, 8(3), 181-199.

[13] Fuest, C. (2018) *"Digitalisation should be promoted, not taxed"*, EconPol Europe Opinion 12, November 2018.

[14] Gaspar, V., P. Collin, M. P. Devereux, J. H. Snabe, T. Varrak, M. Walsh, and B. Westberg (2014) *"Expert group on taxation of the digital economy"*, Brussels: European Commission. Retrived from: http://ec.europa.eu/taxation_customs/business/company-tax/tax-good-governance/expert-group-taxation-digital-economy_en

[15] HMTreasury (2014) *"Finance Bill 2015: Diverted profits tax"*, London: HMRevenue and Custom.

[16] Janeba, E. (1998) *"Tax Competition in Imperfectly Competitive Markets"*, Journal of International Economics, February.

[17] Katz, R. (2015) *"The impact of taxation on the digital economy"*, GSR discussion paper.

[18] Keen, M., and K. A. Konrad (2014) *"The Theory of International Tax Competition and Coordination"*, Max Planck Institute for Tax Law and Public Finance.

[19] Kind, H.J., M. Koethenbuerger, G. and Schjelderup (2008) *"Efficiency enhancing taxation in two-sided markets"*, Journal of Public Economics, 92, 1531-1539.

[20] Kind, H.J., M. Koethenbuerger, and G. Schjelderup (2009) *"On revenue and welfare dominance of ad valorem taxes in two-sided markets"*, Economics Letters, 104(2), 86-88.

[21] Kind, H.J., M. Koethenbuerger, and G. Schjelderup (2010) *"Tax responses in platform industries"*, Oxford Economic Papers, 62(4), 764-783.

[22] Kind, H.J., and M. Koethenbuerger (2017) *"Taxation in digital media markets"*, Journal Public Economic Theory, 00:01-18. https://doi.org/10.1111/jpet.12257

[23] Liberini, F., A. Cuevas, R. Cuevas, A. Lassmann, and A. Russo (2020) *"Taxation and Global Spillovers in the Digital Advertising Market. Theory and Evidence from Facebook"*, CESifo Working Paper, 8149.

[24] OECD (2014) *"Report on the tax challenges of the digital economy"*, Paris: Center for Tax Policy and Administration, Committee on Fiscal Affairs.

[25] OECD (2018) *"Tax Challenges Arising from Digitalisation - Interim Report 2018: Inclusive Framework on BEPS"*, OECD/G20 Base Erosion and Profit Shifting Project, OECD Publishing, Paris.

[26] OECD (2019) *"Addressing the Tax Challenges of the Digitalisation of the Economy"*, Public Consultation Document OECD/G20 Base Erosion and Profit Shifting Project, OECD Publishing, Paris, 13 February - 6 March.

[27] Rochet, J.C., and J. Tirole (2003) *"Platform competition in two-sided markets"*, Journal of the European Economic Association, 1, 990-1029.

[28] Rochet, J.C., and J. Tirole (2006) *"Two-Sided Markets: A Progress Report"*, Rand Journal of Economics, 6, 417-429.

[29] Shaphiro, C. (1980) *"Advertising and Welfare: Comment"*, The Bell Journal of Economics, 11, 749-52.

[30] Spence, M. (1975) *"Monopoly, quality and regulation"*, The Bell Journal of Economics, 417-429.

[31] Stiglitz, J. E. (1982) *"Self-selection and Pareto-efficient taxation"*, Journal of Public Economics, 17 (2), 213-240.

Chapter 3

How privacy disclosure and specific taxes affect a two-sided platform

This paper considers a model where a two-sided digital platform sells services to users and provides targeted ads to a unique retailer to address customers. In the analysis we show how government may use taxes to deal with two kinds of issues: citizens' privacy and fiscal fairness. We use the context of monopoly to exploit the market power of digital platform and an auction to eliminate the problem of double marginalization. Under the assumption that consumers are affected by targeted advertising in their willingness to pay and that they care about their privacy, we show the effects that may arise after taxation or by a change in users' expectation about data mining.

JEL classification: D62, H22, L12, M37.

Keywords: two-sided markets, digital platform, monopoly, taxation, privacy.

3.1 Introduction

Digitalization is undoubtedly the greatest revolution of current century. It has been and continues to be a source of technical and organizational change and one of the main engines of growth, at the point of talking about “digital economy". Multinational tech enterprises are at the top between public companies by market capitalization (PWC, 2019). And prospectives about future growth are high. These big companies are usually multi-products firms, characterized by cross network effects and bilateral market power, as highlighted by Weyl (2010).

Most of them use advertising as one of the revenue sources, but for some of them, ads selling constitutes the only one. In all statistics, internet advertising expenditure has considerably increased in the last decade and nowadays it overcomes all other media. The market presents few ad-selling companies that account for about 75% of the total revenues generated in the advertising market (IAB 2018).

These companies have a considerable market power and in this aspect they resemble monopolies. In fact, they may constitute, if not the only one, at least a preferable channel to address specific bunches of final customers for many retailers. A digital multinational, MNE, usually operates through an interface as a link between two sides of the market. These platforms however usually fail to totally internalize network effects and, as a consequence, they generate distortions (Chiaromonte, 2020).

Policies directed to alleviate these drawbacks therefore must take account of how actions addressed on one side of the market might affect welfare and platform’s behaviour on the other side.

There are further concerns calling for international interventions and homogeneous rules. As highlighted by the European Commission, there exists a discrepancy in fiscal treatment between digital platforms and traditional businesses. Today’s international corporate tax rules do not fit business models that can make profit from digital services in a country without being physically present there. It is not a new phenomenon the fact that digital companies pay much less tax than traditional businesses, even when they make the same profit. This “unfairness" is partially justified by specific aid rules to sustain innovation and R&D, for instance there exist particular laws for start-ups, but, in other cases, it is totally unjustified.

Big digital MNEs benefit from the same advantages and infrastructure as traditional businesses, high speed internet, road, a stable legal system just to name a few, but often they do not contribute at all by paying their “fair" share of taxes in the countries they do business in.

To find a balance between these needs, one that supports innovation, but also to make sure that all companies, digital and traditional, big or small, contribute to society, and operate on the same playing field, is not an easy task. Though it has to be the goal of a new international tax legislation. In doing it, one cannot ignore the consequences on economic players that a change of this dimension would generate. The present paper tries to show some of these aspects, though it focuses on a specific business model, as a simplification of a real one. Firstly, it recognizes the role of data in the value chain. Secondly, it shows some consequences arising from indirect taxes. And last but not least, we pay attention on users' behaviour and their concerns about personal information disclosure.

In the digital economy, value is indeed created, with different weights, from a combination of algorithms, user data, sales functions and know-how. For example, a user can contribute to value creation by sharing his/her preferences (e.g. liking a page, or giving a feedback) on a social platform. This data will later be used and monetised for targeted advertising. The profits however are often taxed where the advertising algorithm has been developed, that it is usually a low tax country. This means on one side that the user contribution to the profits is not considered where the company is taxed, and on the other side that the company can gain an “unfair" fiscal advantage compared to traditional businesses by locating its activities in a low tax countries and acting remotely. Understanding how algorithms work is the key and it is positive that the EU Commission is carrying out an in-depth analysis into algorithmic transparency and other issues through a group of experts (European Commission, 2018).

In Industrial Organization, multi-sided markets have been studied by Anderson and Coate (2015), by Rochet and Tirole (2003, 2006), and Armstrong (2006). Recently, Belleflamme and Toulemonde (2009, 2016) and Belleflamme and Peitz (2019) instead have add intra-group externalities to models with inter-group externalities, and have analysed competition issues, but only marginally fiscal effects. The way taxation impacts on these businesses, focusing specifically on media industry, has been studied by Kind et al. (2008, 2009, 2010). They show how taxation may help to correct some inefficiencies and that the welfare dominance of ad valorem taxes on specific ones does not hold in these markets.

The other aspect related to our paper is users' information. Data are considered in every sector as the "new oil", most of the value created in digital industry come from data, and this oil is essentially free for the platform, since it is provided by users who do not know very well what they are giving away. The true cost, although very consistent in some cases, for the platform is only that of developing and improving the algorithm to read these data. Due to the relevance of data in the chain value and also to tackle privacy issues, some scholars have proposed to levy a tax on persona data (Collin and Colin, 2013). Data taxation has been also investigated by Bourreau et al. (2017) and by Bloch and Demange (2017), to our knowledge.

Our paper studies the impact and the effects of taxing a two-sided monopolistic platform offering personalized services to users and targeted advertising to sellers, based on the collection of users' data. Our goal is to show how consequences may be different according to the taxes used. In particular, we find that the strength of

data externalities determines both the adopted specific business model and how the platform reply to a tax introduction.
The paper is organized as follows: Section 2 presents the model, all the agents involved and how the platform makes its optimal choices; next Section treats what happens following a change in users' expectation about their data exploitation; Section 4 shows the effects of unit taxes on platform's choices; the last Section concludes with some comments and gives some research lines for the future.

3.2 The model

We have a platform which chooses the level of data exploitation s, the amount of ads a to produce and the fee Z which users have to pay to join the platform's services.
The retailer produces a good x, sold to customers on the platform at price p, and participates to an auction paying an amount A equivalent to its profits. Hence, at last the digital platform is able to catch all producer's surplus.
Customers first decide whether to join or not the platform. Joining it, they enjoy a basic service which is a function of the level of their data exploitation. Moreover, they can be addressed through targeted ads by the retailer and as a consequence, they can buy good x. So, they also benefit from the consumption of good x. Finally, they care for their privacy, they know they are giving away their information, but they do not know the actual level of data exploitation.
The game proceeds as follows: at first stage people decide to enter or not the platform, according to the fee Z and on their common expectation about the level of data exploitation s^e, but without the possibility to verify it.
The utility from joining the platform is given by:

$$U = u(s) - \psi(s^e) - Z \tag{3.1}$$

where $u(s)$ is drawn from a uniform distribution on $(0, v)$ with a mass of $v(s)$ of potential users. A user enters the platform if $u(s) \geq \psi(s^e) + Z$. Therefore, denoting by n the number of participants, we have from expression (3.1) that:

$$n = v(s) - \psi(s^e) - Z \tag{3.2}$$

where the utility function $v(s)$ is increasing and concave, $v_s > 0$ and $v_{ss} < 0$, whose subscripts state for the derivative, whereas the disutility function $\psi(s^e)$ is increasing and convex, that is $\psi_{s^e} > 0$ and $\psi_{s^e s^e} > 0$.
Retailer's profits are defined as:

$$\Pi^x = [p(x, a, s)x - c(x)]\, \theta(n) - A \tag{3.3}$$

where $\theta(n)$ indicates the probability per active seller that the users will buy the good x. It is an increasing function in its argument, at no increasing rate, that is $\theta_n > 0$ and $\theta_{nn} \leq 0$. Its first order condition is:

$$p + xp_x = c_x \tag{3.4}$$

Platform's profit function is defined by the following expression:

$$\Pi^p = n * Z + A - k(a, s) \tag{3.5}$$

which can be rewritten as:

$$\Pi^p = [v(s) - \psi(s^e)]\, Z - Z^2 + [p(x, a, s)x - c(x)]\, \theta(n) - k(a, s) \tag{3.6}$$

and whose first order conditions are:

$$Z: \quad v(s) - \psi(s^e) - 2Z - [px - c(x)]\,\theta_n = 0 \tag{3.7}$$

$$a: \quad xp_a\theta(n) = k_a \tag{3.8}$$

$$s: \quad v_s Z + [v(s) - \psi(s^e) - 2Z]\frac{\partial Z}{\partial s} + xp_s n + [px - c(x)]\,\theta_n\left(v_s - \frac{\partial Z}{\partial s}\right) = k_s \tag{3.9}$$

From condition (3.7), one can obtain the equilibrium access fee for users as:

$$Z^* = \frac{v(s) - \psi(s^e) - [px - c(x)]\,\theta_n}{2} \tag{3.10}$$

and from (3.10), one can derive $\frac{\partial Z}{\partial s} = \frac{v_s - xp_s\theta_n}{2}$.

Looking at condition (3.10), when the disutility $\psi(s^e)$ of the expected data exploitation is very high or the earnings on the retailer's side are very large, then the access fee Z could be negative. In this case we assume a corner solution, with $Z = 0$.

Combining the condition (3.10) with (3.2), one can find the number of users joining the platform:

$$n^* = \frac{v(s) - \psi(s^e) + [px - c(x)]\,\theta_n}{2} \tag{3.11}$$

One can find the relation existing between the advertising and the level of data exploitation substituting (3.10) and (3.11), respectively, in (3.8) and in (3.9):

$$\begin{aligned} xp_a\,[v(s) - \psi(s^e) + (px - c(x))\,\theta_n] &= 2k_a \\ v_s\,[v(s) - \psi(s^e) + (px - c(x))\,\theta_n] + 2xp_s\theta(n) &= 2k_s \end{aligned}$$

Or by exploiting condition (3.11), as:

$$xp_a n = k_a \tag{3.12}$$

$$v_s n + xp_s\theta(n) = k_s \tag{3.13}$$

By combining the two conditions above, (3.12) and (3.13), we get the efficient ratio between a^* and s^*:

$$\frac{xp_a}{v_s} = \frac{k_a}{k_s - xp_s\theta(n)} \tag{3.14}$$

For interpretation purposes, consider hereafter $\theta(n) = \lambda n$ with $\lambda = 1$. Hence, expression (3.14) becomes:

$$\frac{xp_a}{v_s + xp_s} = \frac{k_a}{k_s} \tag{3.15}$$

In setting this efficient ratio, the platform aims at balancing the marginal effect on price linked to an increase in the number of ads, p_a, and the double effect of an increase in data exploitation on users' evaluation of the service, v_s and on final price for good x, p_s.

Furthermore, using the simplification about $\theta(n)$, we can rewrite the optimal access fee and derive the number of users as follows:

$$Z^* = \frac{v(s) - \psi(s^e) - px + c(x)}{2} \tag{3.16}$$

$$n^* = \frac{v(s) - \psi(s^e) + px - c(x)}{2} \tag{3.17}$$

So, more profits the platform gains from retailer's side, more likely it will set an access fee equal to zero, greater is the incentive to have more users joining the platform. The reader should remind that the fee A for advertising is exactly equal to the profit made by the retailer.

As a consequence, any action changing this equilibrium would spread its effect on both sides of markets.

In the next section, we will begin by analysing one of these possibilities.

3.3 The effect of a change in users' expectation about their privacy

Firstly, it is interesting to understand what happens to platform's choices in front of a change in privacy-consciousness by users. For instance, an increase in s^e may happen as a consequence to some scandals, such as a disclosure of sensitive information[1], or a more coercive rule about privacy like the introduction of the GDPR[2] in Europe for instance.

What we expect is a twofold effect, both on users' side and on retailer's one. Using comparative statics[3], it is possible to show that both the level of ads a and the level of data exploitation s fall, while, in principle, the access fee Z may either rise or decrease:

$$\begin{aligned}\frac{dZ}{ds^e} &= \frac{(+/-)}{\Omega} \gtrless 0 \\ \frac{da}{ds^e} &= \frac{(+)}{\Omega} < 0 \\ \frac{ds}{ds^e} &= \frac{(+)}{\Omega} < 0\end{aligned}$$

where $\Omega < 0$ since it is the determinant of the negative semidefinite Hessian matrix 3×3 of first order conditions associated to equations (3.7-3.9).

The sign of access fee Z depends strictly on the specific business we are going to analyse and on the relative strengths of externalities. In fact, going through the complex derivation of comparative statics, one can state that higher is the retailer's marginal willingness to pay for targeted advertising, that is p_a and p_s, the more likely Z increases. The platform indeed tries in this way to partially shift the negative effect

[1] An example has been the Cambridge Analytical scandal. See The Economist (2018).

[2] The General Data Protection Regulation (EU) 2016/679 (GDPR) is a regulation in EU law on data protection and privacy in the European Union (EU) and the European Economic Area (EEA). The GDPR aims primarily to give control to individuals over their personal data and to simplify the regulatory environment for international business by unifying the regulation within the EU.

[3] See the Appendix for the whole derivation

of users' expectation by increasing the fee they are going to pay and reducing their participation since the retailer's side is now less profitable. So, also when the access fee is set to zero, consequently to a change in users' expectation it is possible that the platform finds convenient to turn it in a positive value.

On the contrary, when the profitability of retailer's side is not enough to partially compensate for the reduction in targeted advertising, the optimal response is to set a lower access fee to increase the number of users joining the platform.

Hence, to sum up, the final effect can be a partial shift of profits from one side of the market to the other one, to compensate in part the losses, or, on the contrary, a change in expectation may end up by reducing profits twice[4], on both sides of the market.

Proposition 6 *When the externalities of data on users' evaluation of services are low, while their marginal willingness to pay derived from data exploitation is very high, it exists the possibility that the optimal strategy for the platform consists of setting a higher access fee for users and reducing both the number of ads and the level of data exploitation.*

Looking at the consequences on the number of users joining the platform indeed, from totally differentiating $n(s, s^e, Z)$, we get:

$$\frac{dn}{ds^e} = v_s \underset{-}{\frac{ds}{ds^e}} - \psi_{s^e} - \underset{+/-}{\frac{dZ}{ds^e}} \tag{3.18}$$

The condition (3.18) tells us that when the platform finds convenient to increase the access fee, less users will join the platform. Otherwise, when the platform reduces its access fee, it tries to partially compensate the reduction in participants to its service in order to reduce the losses made on the other side of the market. This happens each time the retailer's side is more profitable than users' one.

One can see it better by taking the total differentiation of gross retailer's profit, with

[4]It is similar to what found by Belleflamme and Toulemonde (2016), when they talk about "double jeopardy".

$A = 0$:

$$\frac{d\Pi^x}{ds^e} = x\left[\frac{\partial p}{\partial a}\frac{da}{ds^e} + \frac{\partial p}{\partial s}\frac{ds}{ds^e}\right]n + [px - c(x)]\frac{dn}{ds^e} \tag{3.19}$$

The result depends on the sign of the derivative of n with respect to s^e. The retailer side is less profitable, since the first square bracket sign is negative, but the losses may be partially compensated, at least in principle, by an increase in users' participation, i.e. when $\frac{dn}{ds^e} > 0$. The more likely a business depends on data mining, the more likely the platform will avoid a shrinking in users' number.

3.4 The effects of unit taxes on platform's choices

In line with the current debate about the need to find a way to build a fairer playing field for all businesses without exceptions, single governments are implementing unilateral actions to deal with digital platforms. Their aim is twofold: on one hand, they want to rise fiscal earnings without increasing the fiscal burden on domestic players, on the other hand, they want to fight harmful tax planning made by multinational enterprises, and particularly, by digital businesses, which allows them to avoid any direct form of taxation and to take advantage compared to standard businesses. In order to deal with these two issues, the European Commission is studying some proposals to guarantee a fair level of competition between companies within its borders, but so far no real actions have been implemented due to the lack of a common agreement and the fear of international repercussions. Nevertheless, some countries have adopted unilateral specific fiscal regimes to tax digital companies.
The next discussion aims at highlighting the likely consequences following a unilateral decision made by a single government about adopting specific taxes.

3.4.1 A specif tax t_Z on the access fee

Let us focus on the effects of specific taxation on platform's choices. We use three unit taxes in line with what has already been implemented in some countries and

proposed by Colin and Collin (2013): a specific tax on the access fee, t_Z, one on advertising, t_a, which is measurable in the number of ads displayed online and a specific tax on data exploitation, t_s, measured using data exchange throughout servers as a proxy.
Thus, platform's profit function becomes:

$$\Pi^p = n * Z + A - k(a, s) - t_Z Z - t_a a - t_s s \tag{3.20}$$

and the optimal access fee changes as follows:

$$Z^{**} = \frac{v(s) - \psi(s^e) - px + c(x) - t_Z}{2} \tag{3.21}$$

Looking only at the effects of a tax t_Z, its impact is clearly negative, as confirmed by comparative statics[5]:

$$\begin{aligned} \frac{dZ}{dt_Z} &= \frac{(+)}{\Omega} < 0 \\ \frac{da}{dt_Z} &= \frac{(-)}{\Omega} > 0 \\ \frac{ds}{dt_Z} &= \frac{(-)}{\Omega} > 0 \end{aligned}$$

where $\Omega < 0$, and it is the determinant of the negative semi-definite Hessian matrix 3×3 derived from first order conditions.
As evident the effect of a tax on users' fee is to reduce the profitability for the platform on that side of the market. As a consequence, the monopolist chooses to shift partially the source of its profits from users' to retailer's side, by exploiting more the targeted advertising effect.

Proposition 7 *A specific tax on users' fee changes the profitability of users' side of the market. The platform is induced to partially move the source of its revenues from that side to the retailer's one by reducing the fee and increasing the investment in targeted advertising.*

[5]For the full derivations, see the appendix B.2

At last, the platform sets a lower access fee for users, but it displays more ads and exploits more users' data.

The consequences on the number of users joining the platform are positive. In fact, provided that there is no change in privacy's consciousness, we get:

$$\frac{dn}{dt_Z} = v_s \frac{ds}{dt_Z} - \frac{dZ}{dt_Z} > 0$$

So, the platform finds convenient to exploit the other side of the market by providing better services and reducing taxable profits. As an extreme consequence, the monopolist could set no access fee, $Z = 0$, paying no taxes and making all profits from the retailer's side.

This result is in line with what expected. More interesting is instead looking at two other fiscal instruments and their consequences.

3.4.2 A specific tax on advertising, t_a, or on data, t_s

The effect of a tax on access fee indirectly affects also the optimal ratio in (3.15), since the number of ads a and data s changes, but the optimal rule, which guarantees the efficient provision, still holds.

A unit tax on ads or on data instead distort the optimal ratio:

$$\frac{xp_a}{v_s + xp_s} = \frac{k_a + t_a}{k_s + t_s} \tag{3.22}$$

Only when the increase in numerator and denominator is equally proportional, which is different from saying that $t_a = t_s$, then the ratio does not change, though the quantity of a and s are different compared to the case without taxes.

Moreover, their final effect on the variables chosen by the platform is similar. Both of them may cause either a rise or a decrease in the access fee for users, depending on the strengths of externalities on price of good x, but both a tax on ads, and a tax on data, will likely end by reducing the number of ads displayed and the level of data exploitation.

Studying the effect of a tax on advertising, from comparative statics we get:

$$\frac{dZ}{dt_a} = \frac{(v_s - xp_s)(2nxp_{as} - k_{as}) + xp_a\left[2n(v_{ss} + xp_{ss}) + v_s(v_s + xp_s) - k_{ss}\right]}{\Omega} \gtrless 0$$

$$\frac{da}{dt_a} = \frac{-2\left[2n(v_{ss} + xp_{ss}) + v_s(v_s + xp_s) - k_{ss}\right] + v_s^2 - x^2p_s^2}{\Omega} < 0$$

$$\frac{ds}{dt_a} = \frac{xp_a(v_s + xp_s) + 2(2nxp_{as} - k_{as})}{\Omega} < 0$$

where $\Omega < 0$, and it is the determinant of the negative semi-definite Hessian matrix 3×3 derived from FOCs.

The numerators of the second and third expression are both positive, so the sign is well-defined, and, as a consequence, the final effect on a and s is negative. The first numerator is in fact the principal minor of a matrix 2×2, which is positive when the second order conditions hold. The second one is positive due to the initial assumption of the model about the externalities of a and s on users' utility and willingness to pay.

The sign of the numerator in the first expression is ambiguous. The square bracket is negative and so it pushes for an increase in the access fee Z, whereas the second round bracket is positive, but the first one may be either positive or negative. When it is negative, that means that the externality of s on price weighted by the amount of good x sold is higher than the externality on users' utility, then the access fee rises. Otherwise, when the business is not data-driven, that means that data are more useful for users' services than for creating value for the retailer, then the sign is not well-defined.

For a tax on data, t_s, using comparative statics, we obtain:

$$\frac{dZ}{dt_s} = \frac{-xp_a(2nxp_{as} + xp_a v_s - k_{as}) - (v_s - xp_s)(2nxp_{aa} - k_{aa})}{\Omega} \gtrless 0$$
$$\frac{da}{dt_s} = \frac{-xp_a(v_s - xp_s) + 2(2nxp_{as} + xp_a v_s - k_{as})}{\Omega} \lesseqgtr 0$$
$$\frac{ds}{dt_s} = \frac{-2(2nxp_{aa} - k_{aa}) - x^2 p_a^2}{\Omega} < 0$$

In this case, only the sign of the effect on data is well-defined and negative, since the numerator is a principal minor of a matrix 2×2. Whereas for the first two expressions, the sign depends strongly on the round bracket $(v_s - xp_s)$. When it is negative, the sign is positive for the access fee and negative for the advertising, otherwise it remains ambiguous. Hence, the result is the same as the previous case only for the impact on data exploitation.
The effect is also partially similar in sign to that found for a change in users' evaluation. Further considerations can be made when the data externality is higher for retailer's revenue than for users' benefits, i.e. $xp_s > v_s$.

Proposition 8 *When the externality of data on users' evaluation of services is weaker than the increase of their marginal willingness to pay derived from data exploitation, that is $v_s < xp_s$, then the optimal strategy of platform is well defined. The platform would set a higher access fee for users and would reduce both the number of ads and the level of data exploitation, independently of the implemented unit tax.*

This strategy is justified indeed by the opportunity for the platform partially change the source of its revenue on the other side of the market. In this way, it is able to reduce the fiscal burden on the most profitable side of the market.
Finally, looking at the consequence on the number of users joining the platform, from totally differentiating $n(s, s^e, Z)$, provided that there is no change in privacy's

consciousness, we get:

$$\frac{dn}{dt} = v_s \underset{-}{\frac{ds}{dt_i}} - \underset{+/-}{\frac{dZ}{dt_i}} \tag{3.23}$$

where t_i with $i = a, s$ indicates the implemented specific tax. The condition (3.23) tells us what happens to users' participation.

Proposition 9 *When the externality of data on users' evaluation of services is weaker than the increase of their marginal willingness to pay derived from data exploitation, that is $v_s < xp_s$, then the number of users joining the platform will decrease. The platform would prefer less users' participation to partially compensate the reduction in services' quality and in order to not reduce too much the level of data exploitation.*

So, when the platform finds convenient to increase the access fee, then less users will join the platform. Otherwise, when it reduces its access fee, it tries to partially compensate the reduction in participants caused by poorer services' quality in order to reduce the losses made on the other side of the market due to lower data exploitation. This happens each time the retailer's side is more profitable than users' one. In simple terms, in data driven businesses, profits come from the ability to extract value from personal info. Conversely, in mass driven businesses, profits come mainly from the number of customers and the analysis of their preferences is not so relevant. The strength of data externality determines the adopted specific business model. The more retailer's sales depend on data, the more likely the monopoly would charge a higher access fee. The less the business is data driven, the more likely the monopoly would try to subsidise users' participation.

3.5 Conclusion

In this paper, we deal with a two-sided monopolist platform which makes profit both from selling personal services to users and selling targeted advertising to retailers. We show how this business model tries to balance the profitability of the two sides of the market by choosing an efficient level of data exploitation and setting a non-negative access fee. When the profitability of ads' sales is very high, there exists the convenience for the platform of offering services freely in order to push up users' participation.

We introduce here an element of novelty regarding users' expectation about their privacy. Though it might seem a strong simplification, this aspect allows to show how a change in the attention to personal privacy may have direct effects on platform's choices.

In line with the current debate about the harmful tax planning of digital platform and the need to rise fiscal revenue by the government, the analysis goes on by showing the effects of introducing specific taxes on sources of income for the platform. According to the specific business model adopted by the platform, a data driven business or a mass driven one, the results are opposite. In the former, the monopolist would try to subsidise users' participation by lowering the access fee, conversely in the latter the monopolist would partially shift the source of its income from the retailer to users by charging a higher access fee as best reply.

These results suggest that a policy maker cannot ignore the business model adopted by a digital platform before choosing the best reply to this phenomenon. An empirical enquire is needed to take the best choice and to understand the mechanism behind these companies. In this line, the attention of the EU Commission to study how algorithms work through a group of experts is very positive. Given that most of the business models involved in the digital economy rely on advertising based on data exploitation, we have to draw some actions to balance the needs for personal privacy and to raise fiscal revenue assuring an equal playing field for all businesses, traditional and digital. Single and uncoordinated responses seem to be a poor way for governments to capture a bigger share of the digital value creation, and they may finish by exacerbating distortions, by stifling innovation and by creating a puzzle of fiscal rules. For a global problem, the best solution in our opinion is an international agreement, though it will not be easy to build a common ground among many different countries. Otherwise, the final scenario will be a fragmented landscape, with problems of double taxation and international disputes.

There are several paths to extend the present analysis. One is to investigate how these results change in presence of some form of competition between platforms. Another one consists of studying the effect of data exploitation on investment decisions, how these are made and whether some kind of taxes would finish by stifling innovation or exacerbating the efforts to create value from personal data. Other studies can be performed through experiments to analyse users' perception of their privacy and if and how their privacy-consciousness changes after specific events. However, the greatest effort has to be concentrated into empirical research. In fact, national and supranational institutions may help to force digital companies to disclose their data to scholars and be more transparent in front of citizens.

References

[1] Anderson, P.S., and S. Coate (2005) *"Market provision of broadcasting: a welfare analysis"*, Review of Economic Studies, 72, 947-972.

[2] Armstrong, M. (2006) *"Competition in Two-Sided Markets"*, RAND Journal of Economics, 37, 668-691.

[3] Bacache Beauvallet, M. (2017) *"Tax competition, tax coordination, and e-commerce"*, Journal of Public Economic Theory, 00, 1-18. https://doi.org/10.1111/jpet.12254

[4] Bagwell, K. (2007) *"The Economic Analysis of Advertising"*, Chapter 28 in Handbook of Industrial Organization, vol. 3, pp 1701-1844, Elsevier.

[5] Belleflamme P., and E. Toulemonde (2009) *"Negative Intra-group Externalities in Two-Sided Markets"*, International Economic Review, 50(1), 245-272.

[6] Belleflamme P., and E. Toulemonde (2016) *"Tax Incidence on Competing Two-Sided Platforms: Lucky Break or Double Jeopardy"*, CESifo Working Paper No. 5882.

[7] Belleflamme P., and M. Peitz (2018) *"Inside the engine room of digital platforms: Reviews, ratings, and recommendations"*, Working paper.

[8] Belleflamme P., and M. Peitz (2019) *"Price Disclosure by Two-sided Platforms"*, International Journal of Industrial Organization, 67, 102529.

[9] Bloch F., and G. Demange (2017) *"Taxation and privacy protection on Internet platforms"*, Journal of Public Economic Theory, 00:1-15. https://doi.org/10.1111/jpet.12243

[10] Bourreau M., B. Caillaud, and R. DeNijs (2017) *"Taxation of a digital monopoly platform"*, Journal of Public Economic Theory, 00:1-12. https://doi.org/10.1111/jpet.12255

[11] Chamberlin, E. (1933) *"The Theory of Monopolistic Competition"*, Cambridge, MA: Harvard University Press.

[12] Chiaromonte, N. (2020) *"Targeted Advertising on the web: too little or too much? How taxation might be welfare-enhancing"*, Working paper.

[13] Colin, N. and P. Collin (2013) *"Mission d'expertise sur la fiscalité de l'économie numérique"*, Paris. Retrieved from: http://www.economie.gouv.fr/files/rapport-fiscalite-du-numerique_2013.pdf

[14] Dorfman, R. and P.O. Steiner (1954) *"Optimal Advertising and Optimal Quality"*, America Economic Review, 44, 826-36.

[15] Esayas, S. Y. (2018) *"Competition in (data) privacy:'zero'-price markets, market power, and the role of competition law"*, International Data Privacy Law, 8(3), 181-199.

[16] European Commission (2018) *"Commission Decision of 26.4.2018 on setting up the group of experts for the Observatory on the Online Platform Economy"*, Brussels, 26.4.2018 C(2018) 2393 final.

[17] Gaspar, V., P. Collin, M. P. Devereux, J. H. Snabe, T. Varrak, M. Walsh, and B. Westberg (2014) *"Expert group on taxation of the digital economy"*, Brussels: European Commission. Retrived from:

http://ec.europa.eu/taxation_customs/business/company-tax/tax-good-governance/expert-group-taxation-digital-economy_en

[18] IAB (2018) *"Internet Advertising Revenue Report 2018 Full Year Results"*, PwC.

[19] Katz, R. (2015) *"The impact of taxation on the digital economy"*, GSR discussion paper.

[20] Keen, M., and K. A. Konrad (2014) *"The Theory of International Tax Competition and Coordination"*, Max Planck Institute for Tax Law and Public Finance.

[21] Kind, H.J., M. Koethenbuerger, and G. Schjelderup (2008) *"Efficiency enhancing taxation in two-sided markets"*, Journal of Public Economics, 92, 1531-1539.

[22] Kind, H.J., M. Koethenbuerger, and G. Schjelderup (2009) *"On revenue and welfare dominance of ad valorem taxes in two-sided markets"*, Economics Letters, 104(2), 86-88.

[23] Kind, H.J., M. Koethenbuerger, and G. Schjelderup (2010) *"Tax responses in platform industries"*, Oxford Economic Papers, 62(4), 764-783.

[24] Kind, H.J., and M. Koethenbuerger (2017) *"Taxation in digital media markets"*, Journal Public Economic Theory, 00:01-18. https://doi.org/10.1111/jpet.12257

[25] OECD (2014) *"Report on the tax challenges of the digital economy"*, Paris: Center for Tax Policy and Administration, Committee on Fiscal Affairs.

[26] OECD (2018) *"Tax Challenges Arising from Digitalisation - Interim Report 2018: Inclusive Framework on BEPS"*, OECD/G20 Base Erosion and Profit Shifting Project, OECD Publishing, Paris.

[27] OECD (2019) *"Addressing the Tax Challenges of the Digitalisation of the Economy"*, Public Consultation Document OECD/G20 Base Erosion and Profit Shifting Project, OECD Publishing, Paris, 13 February - 6 March.

[28] PwC (2019) *"Global Top 100 companies by market capitalization"*, available at: https://www.pwc.com/gx/en/audit-services/publications/assets/global-top-100-companies-2019.pdf

[29] Rochet, J.C., and Tirole, J. (2003) *"Platform competition in two-sided markets"*, Journal of the European Economic Association, 1, 990-1029.

[30] Rochet, J.C., and Tirole, J. (2006) *"Two-Sided Markets: A Progress Report"*, Rand Journal of Economics, 6, 417-429.

[31] The Economist (2018) *"Why is Mark Zuckerberg testifying in Congress?"*, April 9th, 2018. Retrieved from: https://www.economist.com/the-economist-explains/2018/04/09/why-is-mark-zuckerberg-testifying-in-congress

Chapter 4

Optimal taxation with persuasive advertising

This paper studies the optimal structure of taxation in a model where persuasive advertising is treated as undesired by the government, which, as consequence, aims at maximising a "laundered" version of consumers' utility. We tackle also distributional concerns, allowing for two types of agents, high-skilled and low-skilled. We find that in some cases a positive tax rate on commodities or on ads can help to pursue redistributive goals. In addition, under some conditions, combining commodity taxation with a positive advertising tax rate can be welfare-enhancing.

JEL classification: H21, D60, M37.

Keywords: Nonlinear income taxation; commodity taxation; redistribution; advertising.

4.1 Introduction

Economists have long recognized that advertising has two main functions: to inform and to persuade.

A third, but less exploited function is associated with the concept of complementarity. Becker and Murphy (1993) was the first to introduce and interpret advertising as a complement of the advertised good.

In the digital age, the information function seems to have been overcome, because consumers can get all the product informations they want from a quick search on web. That makes virtually all advertising today purely persuasive. This is even more true since it seems difficult to compare the whole amount of information available online. As consequence, every day new web-sites are born to help people making right choices. Information function seems to belong to these web-sites than advertising itself.

Persuasive advertising has been considered anti-competitive, because it induces people to buy products that they do not really prefer, harming consumers and placing sellers of consumers' preferred products at a competitive disadvantage.

The idea that some individuals do not know what is best for them is not new in economics and other social sciences. Mistakes in individuals' behaviour is in fact the object of behavioural public economics (Kanbur et al., 2006). Indeed the key feature of behavioural optimal taxation models has been to highlight how standard optimal tax formulas should be amended by the presence of terms reflecting the failure of consumers to act in accordance with their real well-being.

A first attempt at building a link between the empirical literature on the determinants of subjective well-being and the theory has been made by Gerritsen (2016).

While the most recent and complete contribution about optimal taxation with behavioural agents belongs to Farhi and Gabaix (2020). They provide a general model of behavioural biases, which allows on one side to show how the forces arising in isolation interact, and on the other side it adds new insights on some of the cornerstone results of optimal taxation theory.

Although the growing attention in literature to correct “wrong" behaviour by individuals, as highlighted by Thaler and Sunstein (2008), at the best of our knowledge, no one has paid proper attention to distortions in choices induced by advertising.

In the era of digitalization, people are bombed continuously by ads from everywhere at every time, directly or in an inaudible way. This is enough to justify in principle government’s actions to limit and to regulate digital advertising. Furthermore, concerns about privacy and data exploitation in order to create fitter targeted ads might be considered an additional reason to ask authority’s intervention.

The paper conceptually closest to our is Blomquist and Micheletto (2006), which deals with the concepts of paternalism and merit goods, showing the consequences for the structure of direct and indirect taxation of a divergence between the individuals’ and the policy maker’s utility functions.

The way we choose to treat persuasive advertising, such as a demerit good, is also similar to how Veblen effects are treated in Micheletto (2011), who highlights that “laundering" the individuals’ preferences adds a new term in the optimal formulas, which consists of an incentive to under/over-provide the good, according to government’s preferences. Related to these two studies, recently, Aronsson and Johansson-Sterman (2018) have made a comparison among paternalism and Veblen effects. Although the tax motives differ between paternalist and welfarist governments, they find that the policy rules for optimal income taxation may be remarkably similar.

Using a similar approach, here we provide a characterization of an optimal tax system when a government aims at “laundering" consumers’ utility from the effect of persuasive advertising.

In our analysis we also tackle distributional concerns, finding that in some cases a positive tax rate on commodities or, alternatively, on ads can help to pursue redistributive goals. In addition, we also show that combining commodity taxation with a positive advertising tax rate might be welfare enhancing, under some circumstances. These results are achieved under the condition of quasi-linear preferences, which is an assumption that is often made for simplifying purposes in the optimal tax literature. Nevertheless, a more general, though cumbersome, result can be derived without relying on the assumption of quasi-linearity.

Therefore, the contribution of the present work is threefold. It helps to shed light on the age-old link between direct and indirect taxation, studied by Atkinson and Stiglitz (1976). Moreover, it faces an issue not yet discussed in the literature of optimal taxation about the restraint of targeted advertising, when this last could lead to unaware damage for individuals. According to this interpretation, persuasive advertising might be considered undesired as well as demerit goods, such as cigarettes or alcohol and addressed into similar ways[1]. And finally, it introduces the idea that persuasiveness might affect distinct groups of customers differently. Some empirical studies are required to validate this hypothesis.

The structure of the paper is as follows. Section 2 presents the basic model and characterizes the behaviour of agents and firms. In Section 3 we present the government's maximization problem and characterize the properties of an optimal tax system in presence of traditional persuasive advertising. In Section 4 we present a more elaborate model where we also introduce the possibility to use a tax on advertising as an additional policy tool for the government to deal with targeted persuasive advertising. Finally, Section 5 provides concluding remarks and a discussion of possible future extensions of our analysis.

[1]For having an example of how merit goods has been treated into literature, see Racionero (2000, 2001).

4.2 The model

There are two groups of agents, low-skilled and high-skilled. Variables pertaining to low-skilled are denoted by ℓ; variables pertaining to high-skilled are denoted by h. Low-skilled agents are paid a wage rate w^{ℓ} and high-skilled agents are paid a wage rate w^h, with $w^h > w^{\ell}$. The size of total population is normalized to unity and the proportion of low-skilled is denoted by π. Agents can buy two goods: good x_1 is internally produced by perfectly competitive firms and its production price is normalized to one; good x_2 is imported from abroad and is sold by a foreign monopolistic firm. There is no limitation on imports and no duties are charged. This firm can use persuasive advertising, denoted by a, to boost its revenue by increasing the amount that consumers are willing to spend on good x_2. Only the consumption choices of low-skilled agents are affected by persuasive advertising[2], a.
For high-skilled agents preferences are given by

$$U^h = u(x_1) + \phi(x_2) - v(L) \tag{4.1}$$

whereas for low-skilled they are given by

$$U^{\ell} = u(x_1) + \phi(x_2) + \psi(x_2, a) - v(L) \tag{4.2}$$

The functions $u(\cdot)$, $\phi(\cdot)$ are increasing and concave and the function $v(\cdot)$ is increasing and convex. The function $\psi(x_2, a)$ captures the persuasive effect of advertising, which increases the marginal willingness to pay of low-skilled agents for good x_2, and it is assumed to be characterized by the following properties:

$$\frac{\partial \psi(x_2, a)}{\partial x_2} > 0; \frac{\partial \psi(x_2, a)}{\partial a} > 0; \frac{\partial^2 \psi(x_2, a)}{\partial x_2 \partial a} > 0; \frac{\partial^2 \psi(x_2, a)}{\partial x_2 \partial x_2} < 0; \frac{\partial^2 \psi(x_2, a)}{\partial a \partial a} < 0.$$

[2]This assumption will be released in the next sections after having introduced targeted advertising. The idea is that without data, high-skill people are able to protect themselves from persuasive ads at no cost, whereas the advertiser cannot discriminate efficiently among customers.

Notice in particular that the marginal effect on utility of good x_2 is increasing in ads a.

Labor income is denoted by I^j, for $j = \ell, h$, with $I^j = w^j L^j$. The government uses a nonlinear income tax $T(I)$ and a linear commodity tax t on good x_2 to pursue redistributive goals. Moreover, tax policy is also used by the government to correct for the fact that low-skilled agents make suboptimal consumption choices due to the effect of persuasive advertising. The monopolistic foreign firm charges a unitary price p for good x_2; denoting by q the consumer price of this good, we have that $q = p + t$. The game proceeds as follows: the government chooses the fiscal instruments, after that the foreign firm set its price p, then consumers make their consumption and labour choices. Thus, we solve the problem by backward induction in the next paragraph.

4.2.1 Consumers' behavior

Consider first the behavior of high-skilled agents. Defining disposable (i.e. after-tax) income by $B \equiv I - T(I)$ and denoting by $V^h(q, B, I)$ the conditional indirect utility obtained by a high-skilled agent for given values of q, B and I, we have:

$$V^h(q, B, I) = \max_{x_1^h, x_2^h} \left\{ u\left(x_1^h\right) + \phi\left(x_2^h\right) - v\left(\frac{I}{w^h}\right) \mid x_1^h + q x_2^h = B \right\}$$

The first order condition $\phi'\left(x_2^h\right) = q u'\left(x_1^h\right)$, together with the budget constraint $x_1^h + q x_2^h = B$ define the conditional demand functions

$$x_i^h = x_i^h(q, B) \quad \text{for } i = 1, 2 \tag{4.3}$$

High-skilled agents choose labor supply maximizing $V^h(q, B, I)$ subject to the link between pre-tax income and after-tax income implied by the income tax schedule $T(I) = I - B$. This allows to implicitly define the marginal income tax rate faced by a high-skilled agent as

$$T'(I) = 1 + \frac{\partial V^h / \partial I}{\partial V^h / \partial B} = 1 - MRS_{IB}^h$$

where MRS_{IB}^h denotes the marginal rate of substitution between I and B for an agent of type h.

Similarly, denoting by $V^\ell(q, B, I; a)$ the conditional indirect utility obtained by a low-skilled agent for given values of q, B, I and a, we have:

$$V^\ell(q, B, I; a) = \max_{x_1^\ell, x_2^\ell} \left\{ u\left(x_1^\ell\right) + \phi\left(x_2^\ell\right) + \psi\left(x_2^\ell, a\right) - v\left(\frac{I}{w^\ell}\right) \mid x_1^\ell + q x_2^\ell = B \right\}$$

The first order condition $\phi'\left(x_2^\ell\right) + \frac{\partial \psi\left(x_2^\ell, a\right)}{\partial x_2^\ell} = q u'\left(x_1^\ell\right)$, together with the budget constraint $x_1^\ell + q x_2^\ell = B$ define the conditional demand functions

$$x_i^\ell = x_i^\ell(q, B; a) \quad \text{for } i = 1, 2. \tag{4.4}$$

Low-skilled agents choose labor supply maximizing $V^\ell(q, B, I)$ subject to the link between pre-tax income and after-tax income implied by the income tax schedule $T(I)$.

This allows to implicitly define the marginal income tax rate faced by a low-skilled agent as

$$T'(I) = 1 + \frac{\partial V^\ell / \partial I}{\partial V^\ell / \partial B} = 1 - MRS_{IB}^\ell$$

where MRS_{IB}^ℓ denotes the marginal rate of substitution between I and B for an agent of type ℓ. Notice that, given the separability structure that characterizes the individuals' preferences, it follows that consumption choices (4.3)-(4.4) do not directly depend on labor supply I/w.

4.2.2 Foreign firm maximization problem

Denoting by $\theta(a)$ the increasing and convex function describing the cost incurred for advertising purposes and assuming that good x_2 can be produced at a constant marginal cost c, the foreign firm solves the following maximization problem:

$$\max_{p,a} \left[\pi x_2^{\ell}\left(B^{\ell}, a, p+t\right) + (1-\pi)\, x_2^{h}\left(B^{h}, p+t\right)\right](p-c) - \theta(a)$$

The first order conditions are:

$$\pi x_2^{\ell} + (1-\pi)\, x_2^{h} + \left[\pi \frac{\partial x_2^{\ell}}{\partial q} + (1-\pi)\frac{\partial x_2^{h}}{\partial q}\right](p-c) = 0 \tag{4.5}$$

$$(p-c)\,\pi \frac{\partial x_2^{\ell}}{\partial a} - \theta'(a) = 0 \tag{4.6}$$

Totally differentiating the set of first order conditions above, and assuming that the second order conditions for a maximum are satisfied, i.e. that the Hessian matrix is negative semi-definite so that its determinant Ω is positive,[3] we get the following comparative statics results[4]:

$$\begin{aligned}\Omega\frac{dp}{dt} &= \pi^2 (p-c)\frac{\partial^2 x_2^{\ell}}{\partial q \partial a}\left[\frac{\partial x_2^{\ell}}{\partial a} + (p-c)\frac{\partial^2 x_2^{\ell}}{\partial q\partial a}\right] \\ &- \left[\pi\frac{\partial^2 x_2^{\ell}}{\partial a\partial a}(p-c) - \theta''(a)\right]\left[\pi\frac{\partial x_2^{\ell}}{\partial q} + (1-\pi)\frac{\partial x_2^{h}}{\partial q}\right] \\ &- \left[\pi\frac{\partial^2 x_2^{\ell}}{\partial a\partial a}(p-c) - \theta''(a)\right]\left[\pi\frac{\partial^2 x_2^{\ell}}{\partial q\partial q} + (1-\pi)\frac{\partial^2 x_2^{h}}{\partial q\partial q}\right](p-c)\end{aligned} \tag{4.7}$$

[3]The determinant Ω is given by:

$$\begin{aligned}\Omega &= \left\{\left[\pi\frac{\partial^2 x_2^{\ell}}{\partial q\partial q} + (1-\pi)\frac{\partial^2 x_2^{h}}{\partial q\partial q}\right](p-c) + \left(\pi\frac{\partial x_2^{\ell}}{\partial q} + (1-\pi)\frac{\partial x_2^{h}}{\partial q}\right)2\right\}\left[\pi\frac{\partial^2 x_2^{\ell}}{\partial a\partial a}(p-c) - \theta''(a)\right] \\ &-\pi^2\left[\frac{\partial x_2^{\ell}}{\partial a} + (p-c)\frac{\partial^2 x_2^{\ell}}{\partial q\partial a}\right]^2\end{aligned}$$

[4]See Appendix C.1 for the whole derivation.

$$\Omega \frac{da}{dt} = \left[\pi \frac{\partial x_2^\ell}{\partial q} + (1-\pi) \frac{\partial x_2^h}{\partial q} \right] \left[\frac{\partial x_2^\ell}{\partial a} - (p-c) \frac{\partial^2 x_2^\ell}{\partial q \partial a} \right] \pi + (p-c) \frac{\partial x_2^\ell}{\partial a} \left[\pi \frac{\partial^2 x_2^\ell}{\partial q \partial q} + (1-\pi) \frac{\partial^2 x_2^h}{\partial q \partial q} \right] \pi < 0 \tag{4.8}$$

$$\Omega \frac{dp}{dB^h} = -(1-\pi) \left[\frac{\partial x_2^h}{\partial B^h} + (p-c) \frac{\partial^2 x_2^h}{\partial B^h \partial q} \right] \left[\pi \frac{\partial^2 x_2^\ell}{\partial a \partial a} (p-c) - \theta''(a) \right] \geq 0 \tag{4.9}$$

$$\Omega \frac{da}{dB^h} = \pi (1-\pi) \left[\frac{\partial x_2^h}{\partial B^h} + (p-c) \frac{\partial^2 x_2^h}{\partial B^h \partial q} \right] \left[\frac{\partial x_2^\ell}{\partial a} + (p-c) \frac{\partial^2 x_2^\ell}{\partial q \partial a} \right] \geq 0 \tag{4.10}$$

$$\Omega \frac{dp}{dB^\ell} = \pi^2 (p-c) \frac{\partial^2 x_2^\ell}{\partial B^\ell \partial a} \left[\frac{\partial x_2^\ell}{\partial a} + (p-c) \frac{\partial^2 x_2^\ell}{\partial q \partial a} \right] - \pi \left[\frac{\partial x_2^\ell}{\partial B^\ell} + (p-c) \frac{\partial^2 x_2^\ell}{\partial B^\ell \partial q} \right] \left[\pi \frac{\partial^2 x_2^\ell}{\partial a \partial a} (p-c) - \theta''(a) \right] \geq 0 \tag{4.11}$$

$$\Omega \frac{da}{dB^\ell} = \pi^2 \left[\frac{\partial x_2^\ell}{\partial a} + (p-c) \frac{\partial^2 x_2^\ell}{\partial q \partial a} \right] \left[\frac{\partial x_2^\ell}{\partial B^\ell} + (p-c) \frac{\partial^2 x_2^\ell}{\partial B^\ell \partial q} \right] - 2\pi (p-c) \frac{\partial^2 x_2^\ell}{\partial B^\ell \partial a} \left[\pi \frac{\partial x_2^\ell}{\partial q} + (1-\pi) \frac{\partial x_2^h}{\partial q} \right] - \pi (p-c)^2 \frac{\partial^2 x_2^\ell}{\partial B^\ell \partial a} \left[\pi \frac{\partial^2 x_2^\ell}{\partial q \partial q} + (1-\pi) \frac{\partial^2 x_2^h}{\partial q \partial q} \right] \geq 0 \tag{4.12}$$

The only expression whose sign is not well-defined is the one capturing the impact of commodity tax t on price p, since on the right hand side of (4.7) the first row is positive, whereas the last two lines are negative. As expected the impact of the tax on ads is negative since advertising is complement to consumption of the advertised good. An increase in either B^ℓ or B^h rises both the price p and the amount of ads.

4.3 The government's problem

From the government's perspective the effect of advertising is harmful since it distorts the behaviour of low-skilled agents who are induced to act as consumers in a way that is not consistent with the maximization of their true well-being. Therefore, the government's optimal tax problem can be formally described as follows:

$$\max_{I^\ell, B^\ell, I^h, B^h, t} \quad V^\ell \left(p+t, B^\ell, I^\ell; a\right) - \psi\left(x_2^\ell, a\right)$$

subject to:

$$\begin{aligned} &V^h\left(p+t, B^h, I^h\right) \geq \overline{V}, \\ &V^h\left(p+t, B^h, I^h\right) \geq V^h\left(p+t, B^\ell, I^\ell\right), \\ &\left(I^\ell - B^\ell + t x_2^\ell\right)\pi + \left(I^h - B^h + t x_2^h\right)(1-\pi) \geq \overline{R}. \end{aligned}$$

In the problem above the first constraint prescribes a minimum utility that should be granted to high-skilled agents; the second constraint is a self-selection constraint requiring high-skilled agents not to mimic low-skilled agents; the final constraint represents the government's budget constraint, where $\overline{R}$ represents an exogenous revenue requirement.

We implicitly assume that the government aims at redistributing towards low-skilled agents so that we do not need to worry about the possibility that low-skilled agents may be tempted to mimic high-skilled agents.

Notice that, rather than aiming at maximizing the utility function of low-skilled agents (subject to the relevant constraints), the government's goal is to maximize a "laundered" version of the utility of low-skilled. In particular, and in accordance with the interpretation of advertising as persuasive rather than informative, the effect of ads is neglected in the "laundered" version of the low-skilled utility.

4.3.1 Optimal commodity taxation

Denote the Lagrange multipliers associated to the three constraints above by, respectively, δ, λ and μ. Denote by a "hat" a variable pertaining to a high-skilled behaving as a mimicker. Denote Hicksian (compensated) demands by a "tilde" symbol. Using first order conditions, after some calculations, it is possible to obtain a complex expression for optimal commodity tax rate[5].

To simplify the interpretation of the optimality condition, assume that the agents' utility function is quasi-linear[6] in x_1 so that $\partial x_2^j/\partial q = \partial \widetilde{x}_2^j/\partial q$ for $j = \ell, h$, $da/dB^\ell = da/dB^h = dp/dB^\ell = dp/dB^h = 0$, and $\partial V^\ell/\partial B^\ell = \partial V^h/\partial B^h = \partial \widehat{V}/\partial B^\ell = 1$. Then, the following Proposition 10 characterizes the optimal commodity tax rate.

Proposition 10 *Under the assumption of quasi-linear utility function in x_1, the expression for optimal commodity tax rate provided by a monopoly becomes:*

$$\begin{aligned} t \;=\; & \frac{\lambda}{\mu Y}\left(x_2^\ell - \widehat{x}_2\right)\left(1+\frac{dp}{dt}\right) \\ & + \frac{1}{\mu Y}\left[\frac{\partial x_2^\ell}{\partial a}\frac{da}{dt} + \frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)\right]\frac{\partial \psi}{\partial x_2^\ell} \\ & - \frac{p-c}{Y}\left[\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q}\right]\frac{dp}{dt} \end{aligned} \tag{4.13}$$

where

$$1+\frac{dp}{dt} = \frac{\left(\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q}\right)\left[\pi\frac{\partial^2 x_2^\ell}{\partial a \partial a}(p-c) - \theta''(a)\right] - \pi^2\left[\frac{\partial x_2^\ell}{\partial a} + \frac{\partial^2 x_2^\ell}{\partial q \partial a}(p-c)\right]\frac{\partial x_2^\ell}{\partial a}}{\Omega}.$$

[5]See the Appendix C.2 for the whole derivation.

[6]It worth to be noted that this assumption is not without consequences. Doing it, we rule out any income effect from the analysis. There exist two different ways of thinking about preferences to be quasi-linear. It is usually assumed that if consumer has to make choices among large variety of goods, then his preferences are quasi-linear at least in one of them. And it constitutes a first meaning of good x_1. Otherwise, quasi-linearity in money is usually assumed, since money is a natural numeraire in which one can express the value of every good. According to this last interpretation, x_1, can be reread as money instead of a good produced in a perfectly competitive market.

and Y *has been defined as*

$$\Upsilon \equiv \left[\pi \frac{\partial \widetilde{x}_2^\ell}{\partial q} + (1-\pi)\frac{\partial \widetilde{x}_2^h}{\partial q}\right]\left(1+\frac{dp}{dt}\right) + \pi \frac{\partial x_2^\ell}{\partial a}\frac{da}{dt}$$

Proof. See Appendix C.2. ■

Noticing that $1 + dp/dt = dq/dt$, i.e. it represents the net variation in the consumer price for good x_2 when t is marginally increased. We will hereafter base our discussion of (4.13) on the assumption that $1 + dp/dt > 0$.

The first term on the right hand side of (4.13) tells us how self-selection considerations affect the choice of the optimal value for t. We know that $x_2^\ell - \widehat{x}_2 > 0$ as persuasive advertising makes low-skilled agents spend more income on good x_2 than a high-skilled mimicker; therefore, since $dq/dt > 0 \Longrightarrow \Upsilon < 0$, self-selection considerations call for subsidizing the consumption of good x_2. The second term on the right hand side of (4.13) is instead a corrective term that descends from the fact that the government launders the preferences of low-skilled agents in the objective function that it maximizes. Given that $da/dt < 0$ (see (4.8)), this term would favour selecting a positive tax rate on the consumption of good x_2. Finally, the last term on the right hand side of (4.13) is a corrective term that has opposite sign to dp/dt. If, as one were to expect, $dp/dt < 0$, the last term on the right hand side of (4.13) would call for setting $t > 0$. This would induce the foreign monopolist to set its price closer to the marginal cost c, therefore reducing the deadweight loss descending from the fact that, exploiting its market power, the monopolist charges a price which is higher than the marginal cost.

To provide a better intuition for this corrective term appearing in (4.13), consider for illustrative purposes the hypothetical case when advertising is not a choice variable for the foreign monopolist, but is instead exogenously given. Then (4.13) would reduce to:

$$t = \frac{\lambda\left(x_2^\ell - \widehat{x}_2\right) + \frac{\partial x_2^\ell}{\partial q}\frac{\partial \psi}{\partial x_2^\ell}}{\mu\left[\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q}\right]} - (p-c)\frac{\frac{dp}{dt}}{1+\frac{dp}{dt}}$$

Hence, if advertising were treated instead as exogenously given, we would have

$$\frac{dp}{dt} = -\frac{\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q} + (p-c)\,\pi\frac{\partial^2 x_2^\ell}{\partial q\partial q} + (p-c)(1-\pi)\frac{\partial^2 x_2^h}{\partial q\partial q}}{\left\{\left[\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q}\right]2 + (p-c)\,\pi\frac{\partial^2 x_2^\ell}{\partial q\partial q} + (p-c)(1-\pi)\frac{\partial^2 x_2^h}{\partial q\partial q}\right\}} < 0$$

so that

$$1+\frac{dp}{dt} = \frac{\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q}}{\left\{\left[\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q}\right]2 + (p-c)\,\pi\frac{\partial^2 x_2^\ell}{\partial q\partial q} + (p-c)(1-\pi)\frac{\partial^2 x_2^h}{\partial q\partial q}\right\}} > 0$$

and therefore:

$$\begin{aligned}\frac{\frac{dp}{dt}}{1+\frac{dp}{dt}} &= \frac{dp/dt}{dq/dt} = -\frac{\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q} + (p-c)\,\pi\frac{\partial^2 x_2^\ell}{\partial q\partial q} + (p-c)(1-\pi)\frac{\partial^2 x_2^h}{\partial q\partial q}}{\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q}} \\ &= -\left[1 + \frac{\pi\frac{\partial^2 x_2^\ell}{\partial q\partial q} + (1-\pi)\frac{\partial^2 x_2^h}{\partial q\partial q}}{\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q}}(p-c)\right] < 0\end{aligned}$$

Thus, when advertising is not a choice variable for the foreign monopolist, the corrective terms for “laundering" and “distortion" call undoubtedly for a positive tax rate.

4.3.2 Optimal marginal income tax rates and marginal effective tax rates

Here we show firstly how to derive the optimal marginal income tax rates for both agents and then their marginal effective tax rates.

For the high skilled we have:

$$T'\left(I^h\right) = 1 - MRS^h_{IB} = 1 + \frac{\frac{\partial V^h}{\partial I^h}}{\frac{\partial V^h}{\partial B^h}}$$

The first order condition with respect to I^h is:

$$(\delta+\lambda)\frac{\partial V^h}{\partial I^h} = -\mu(1-\pi)$$

Whereas, the first order condition with respect to B^h is:

$$(\delta+\lambda)\frac{\partial V^h}{\partial B^h} = -\mu\left[t\frac{\partial x_2^h}{\partial B^h} - 1\right](1-\pi) - \left[\mu t\pi - \frac{\partial\psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dB^h} - \Delta\frac{dp}{dB^h}$$

where Δ accounts for the partial effect of a change in p on the government maximization problem[7].

Combining the two first order conditions above gives:

$$\frac{\frac{\partial V^h}{\partial I^h}}{\frac{\partial V^h}{\partial B^h}}\left\{-\mu\left[t\frac{\partial x_2^h}{\partial B^h} - 1\right](1-\pi) - \left[\mu t\pi - \frac{\partial\psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dB^h} - \Delta\frac{dp}{dB^h}\right\} = -\mu(1-\pi)$$

and rearranging:

$$1 + \frac{\frac{\partial V^h}{\partial I^h}}{\frac{\partial V^h}{\partial B^h}} = T'\left(I^h\right) = \frac{\frac{\partial V^h}{\partial I^h}}{\frac{\partial V^h}{\partial B^h}}\left\{t\frac{\partial x_2^h}{\partial B^h} + \frac{\left[\mu t\pi - \frac{\partial\psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dB^h} + \Delta\frac{dp}{dB^h}}{\mu(1-\pi)}\right\}$$

Assuming that utility is linear in x_1 implies that the expression above simplifies to:

$$T'\left(I^h\right) = 0$$

Therefore, defining the marginal effective tax rate faced by high-skilled agents as

$$METR^h \equiv T'\left(I^h\right) + t\left[\frac{\partial x_2^h}{\partial I^h} + MRS^h_{IB}\frac{\partial x_2^h}{\partial B^h}\right]$$

[7]See equation (C.1) in the Appendix.

we also have that $METR^h = 0$ since $\frac{\partial x_2^h}{\partial I^h} = \frac{\partial x_2^h}{\partial B^h} = 0$[8].

For the low-skilled, instead, we have:

$$T'\left(I^\ell\right) = 1 - MRS_{IB}^\ell = 1 + \frac{\frac{\partial V^\ell}{\partial I^\ell}}{\frac{\partial V^\ell}{\partial B^\ell}}$$

The first order condition with respect to I^ℓ is:

$$\frac{\partial V^\ell}{\partial I^\ell} = \lambda \frac{\partial \widehat{V}}{\partial I^\ell} - \mu\pi$$

The first order condition with respect to B^ℓ is:

$$\frac{\partial V^\ell}{\partial B^\ell} = \frac{\partial \psi}{\partial x_2^\ell}\frac{\partial x_2^\ell}{\partial B^\ell} + \lambda\frac{\partial \widehat{V}}{\partial B^\ell} - \mu\left[t\frac{\partial x_2^\ell}{\partial B^\ell} - 1\right]\pi - \left[\mu t\pi - \frac{\partial \psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dB^\ell} - \Delta\frac{dp}{dB^\ell}$$

Combining the two first order conditions above gives:

$$\frac{\frac{\partial V^\ell}{\partial I^\ell}}{\frac{\partial V^\ell}{\partial B^\ell}}\left\{\frac{\partial \psi}{\partial x_2^\ell}\frac{\partial x_2^\ell}{\partial B^\ell} + \lambda\frac{\partial \widehat{V}}{\partial B^\ell} - \mu\left[t\frac{\partial x_2^\ell}{\partial B^\ell} - 1\right]\pi - \left[\mu t\pi - \frac{\partial \psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dB^\ell} - \Delta\frac{dp}{dB^\ell}\right\} = \lambda\frac{\partial \widehat{V}}{\partial I^\ell} - \mu\pi$$

and rearranging:

$$\mu\pi\left[1 + \frac{\frac{\partial V^\ell}{\partial I^\ell}}{\frac{\partial V^\ell}{\partial B^\ell}}\right] = \lambda\frac{\partial \widehat{V}}{\partial I^\ell} - \frac{\frac{\partial V^\ell}{\partial I^\ell}}{\frac{\partial V^\ell}{\partial B^\ell}}\left\{\frac{\partial \psi}{\partial x_2^\ell}\frac{\partial x_2^\ell}{\partial B^\ell} + \lambda\frac{\partial \widehat{V}}{\partial B^\ell} - \mu t\frac{\partial x_2^\ell}{\partial B^\ell}\pi - \left[\mu t\pi - \frac{\partial \psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dB^\ell} - \Delta\frac{dp}{dB^\ell}\right\}$$

Hence, we get the marginal income tax for low-skilled agents as:

$$\begin{aligned} T'\left(I^\ell\right) \quad &= \quad \frac{\lambda}{\mu\pi}\frac{\partial \widehat{V}}{\partial B^\ell}\left(\frac{\frac{\partial \widehat{V}}{\partial I^\ell}}{\frac{\partial \widehat{V}}{\partial B^\ell}} - \frac{\frac{\partial V^\ell}{\partial I^\ell}}{\frac{\partial V^\ell}{\partial B^\ell}}\right) + t\frac{\partial x_2^\ell}{\partial B^\ell}\frac{\frac{\partial V^\ell}{\partial I^\ell}}{\frac{\partial V^\ell}{\partial B^\ell}} \\ &\quad -\frac{1}{\mu\pi}\frac{\frac{\partial V^\ell}{\partial I^\ell}}{\frac{\partial V^\ell}{\partial B^\ell}}\left\{\frac{\partial \psi}{\partial x_2^\ell}\frac{\partial x_2^\ell}{\partial B^\ell} - \left[\mu t\pi - \frac{\partial \psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dB^\ell} - \Delta\frac{dp}{dB^\ell}\right\} \end{aligned}$$

[8]Therefore the "no distortion at the top" result is not violated, but it strongly depends on the quasi-linearity assumption for preferences.

Assuming that utility is linear in x_1 implies that the expression above simplifies to[9]:

$$T'\left(I^{\ell}\right) = \frac{\lambda}{\mu\pi}\left(\frac{\partial\widehat{V}}{\partial I^{\ell}} - \frac{\partial V^{\ell}}{\partial I^{\ell}}\right) > 0$$

Furthermore, defining the marginal effective tax rate faced by low-skilled agents as

$$METR^{\ell} \equiv T'\left(I^{\ell}\right) + t\left[\frac{\partial x_2^{\ell}}{\partial I^{\ell}} + MRS_{IB}^{\ell}\frac{\partial x_2^{\ell}}{\partial B^{\ell}}\right]$$

we have that $METR^{\ell} \equiv T'\left(I^{\ell}\right)$ since $\frac{\partial x_2^{\ell}}{\partial I^{\ell}} = \frac{\partial x_2^{\ell}}{\partial B^{\ell}} = 0$.
Notice that the impact on $METR^{i}$ of commodity taxation is different from zero when the quasi-linearity assumption is released.

4.4 A model with data and advertising

In this section we try to sophisticate the previous model and to adapt it to digital economy, specifically to targeted advertising industry, like social platforms for instance. By doing it, we consider the use of data as a tool to address customers through targeted ads.
First of all, we need to adapt our previous model to standard digital businesses. Assume that the monopolist selling good x_2 is now a domestic firm and assume that its profits are fully taxed away by the government. Assume also that this monopolistic firm buys advertising from a platform which is located in a foreign country[10]. Moreover, differently from the previous model, the platform is able to recognize both agents and discriminate among them by using targeted ads through personal data exploitation.

[9]The marginal tax rate for low-skilled can be written also as: $T'\left(I^{\ell}\right) = \frac{\lambda}{\mu\pi}\frac{\partial\widehat{V}}{\partial B^{\ell}}\left(MRS_{IB}^{\ell}\right) - \widehat{MRS}_{IB}$; which is necessarily positive when utility is linear in x_1. The standard single-crossing condition. which guarantees that the expression in round bracket is positive, is no longer necessarily satisfied if preferences are not linear into x_1.

[10]In this way, the real profits arising from advertising cannot be addressed by a standard corporate income tax.

Hence, a new element, data s, enters as a proxy for privacy. Implicitly, we are assuming that individuals are not able to evaluate privacy protection issues, unlike government does[11].

Our goal here it is not to make a comparison between the previous model and a new one, but it consists of adapting that one to deal with digital targeted advertising. Therefore, assume high- and low-skilled preferences are represented respectively by the quasi-linear functions:

$$U^h = x_1 + \phi(x_2) + \rho^h \psi(x_2, a, s) - v(L)$$

$$U^\ell = x_1 + \phi(x_2) + \rho^\ell \psi(x_2, a, s) - v(L)$$

where s represents data collected by the platform which gives access to the users at no charge, and ρ^ℓ and ρ^h are two positive constants capturing the extent to which advertising distorts the behavior of, respectively, low- and high-skilled agents[12].

The sequence of game does not change. The government chooses the taxation, after the platform sets the price for ads and sells them to the domestic firm. Then the firm decides the price for commodity x_2 and finally people choose how much to work and how to allocate their earnings. So we can solve it by backward induction.

[11]Though it seems a strong assumption, the concerns about the use of personal information are on the basis of current General Data Protection Regulation (GDPR) 2016/679 in the European Union. These concerns also justify in some way our presumption about the greater attention by countries' authorities rather than a single person.

[12]It seems unrealistic to assume that high-skilled agents are immune (as done before) from ads when these are well-targeted.

The maximization problem solved by the monopolist selling good x_2 is the following[13]:

$$\max_{p,a^h,a^\ell} \quad (1-\tau)\left\{\left[\pi x_2^\ell\left(p+t,a^\ell,s\right)+(1-\pi)\,x_2^h\left(p+t,a^h,s\right)\right](p-c)-\left(a^h+a^\ell\right)p_a\right\}$$

whose first order conditions are:

$$\begin{aligned}
\pi x_2^\ell+(1-\pi)\,x_2^h+\left[\pi\frac{\partial x_2^\ell}{\partial q}+(1-\pi)\,\frac{\partial x_2^h}{\partial q}\right](p-c) &= 0\\
(p-c)\,(1-\pi)\,\frac{\partial x_2^h}{\partial a^h}-p_a &= 0\\
(p-c)\,\pi\frac{\partial x_2^\ell}{\partial a^\ell}-p_a &= 0
\end{aligned}$$

Totally differentiating the three first order conditions above allows deriving expressions for:

$$\begin{aligned}
\frac{dp}{dt} &= \frac{dp}{dt}\,\Big|_{\substack{ds=0\\dp_a=0}}+\frac{dp}{dp_a}\frac{dp_a}{dt}+\frac{dp}{ds}\frac{ds}{dt}\\
\frac{da^h}{dt} &= \frac{da^h}{dt}\,\Big|_{\substack{ds=0\\dp_a=0}}+\frac{da^h}{dp_a}\frac{dp_a}{dt}+\frac{da^h}{ds}\frac{ds}{dt}\\
\frac{da^\ell}{dt} &= \frac{da^\ell}{dt}\,\Big|_{\substack{ds=0\\dp_a=0}}+\frac{da^\ell}{dp_a}\frac{dp_a}{dt}+\frac{da^\ell}{ds}\frac{ds}{dt}
\end{aligned}$$

Assuming that the advertising is produced by the platform at a cost given by the increasing and convex function $g\left(a^h+a^\ell,s\right)$, the maximization problem solved by the platform is given by:

$$\max_{p_a,s} \quad \left(a^h+a^\ell\right)p_a-g\left(a^h+a^\ell,s\right)$$

[13]It should be noted that the prices are linear though the possibility of discrimination among individuals happens in reality. We are aware that it constitutes a limitation of the current approach, but it could be justified in this specific framework. Indeed, having an homogeneous good, it is possible to have some arbitrage between buyers of different skills if we allow for distinct prices without avoiding people to exchange goods.

Here, we implicitly assume that the platform does not discriminate between the charged price for targeted advertising addressed to high- and low-skilled customers[14]. The first order conditions are:

$$\begin{aligned} a^h + a^\ell + \left(\frac{\partial a^h}{\partial p_a} + \frac{\partial a^\ell}{\partial p_a}\right)(p_a - g_1') &= 0 \\ \left(\frac{\partial a^h}{\partial s} + \frac{\partial a^\ell}{\partial s}\right)(p_a - g_1') - g_2' &= 0 \end{aligned}$$

where g_1' and g_2' are respectively the first derivative of function $g(\cdot)$ with respect to the first and the second argument.

Totally differentiating the two first order conditions above one gets expression for dp_a/dt and ds/dt, whose sign is not well-defined[15].

The government's problem is

$$\max_{I^\ell, B^\ell, I^h, B^h, t} V^\ell\left(p + t, B^\ell, I^\ell; a, s\right) - \rho^\ell \psi\left(x_2^\ell, a, s\right)$$

subject to:

$$\begin{aligned} &V^h\left(p + t, B^h, I^h; a, s\right) - \rho^h \psi\left(x_2^h, a, s\right) \geq \overline{V}, \\ &V^h\left(p + t, B^h, I^h; a, s\right) \geq V^h\left(p + t, B^\ell, I^\ell; a, s\right), \\ &\left(I^\ell - B^\ell + t x_2^\ell\right)\pi + \left(I^h - B^h + t x_2^h\right)(1 - \pi) + \\ &+\tau\left\{\left[\pi x_2^\ell\left(p + t, a^\ell, s\right) + (1 - \pi)\, x_2^h\left(p + t, a^h, s\right)\right](p - c) - \left(a^h + a^\ell\right) p_a\right\} \geq \overline{R}. \end{aligned}$$

Once again, the first constraint prescribes a minimum utility that should be granted to high-skilled agents; the second constraint is a self-selection constraint requiring high-skilled agents not to mimic low-skilled agents; the final constraint represents

[14] We are aware of this strong assumption, and that it is not completely true. In fact, it happens in practice to pay for the targeted audience. A firm does it through a complex system of auctions (i.e., a first-price sealed-bid auction for instance). We are going to face this issue in a next paper.

[15] See Appendix C.3 for the relative comparative statics.

the government's budget constraint, where $\overline{R}$ represents an exogenous revenue requirement.

Notice that the second line of the government's budget constraint represents the revenue collected through the tax, levied at rate τ, on the profits of the domestic firm selling good x_2. Since we assume that these profits can be fully taxed by the government[16], in our calculations below we will implicitly assume that $\tau = 100\%$.

The first order conditions of the government's problem with respect to B^h and B^ℓ are, respectively (remember that we have assumed that utility is quasi-linear in good x_1):

$$(\delta + \lambda) - \mu (1 - \pi) = 0 \tag{4.14}$$

$$1 - \lambda - \mu\pi = 0 \tag{4.15}$$

Assume that the nonlinear income tax is optimally chosen so that (4.14)-(4.15) are satisfied. Then consider a tax reform that marginally increases t while at the same time adjusts B^ℓ and B^h by, respectively, $dB^\ell = x_2^\ell dt$ and $dB^h = x_2^h dt$. By considering the effects on the Lagrangian of the government's problem, we can use a perturbation method to derive an expression characterizing the optimal commodity tax rate. As shown in the Appendix, this is given by the Proposition 11:

[16] Without this assumption, we should discuss how to share firm's ownership between high- and low-skilled agents.

Proposition 11 *When t is optimally chosen, the optimal tax rate on commodity provided by a monopolist will be:*

$$\begin{aligned} t = {} & \frac{\lambda}{\mu\Xi}\left(x_2^\ell - \widehat{x}_2\right)\left(1+\frac{dp}{dt}\right) \\ & + \frac{\rho^\ell}{\mu\Xi}\frac{\partial\psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right) + \frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt} + \frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right] \\ & + \frac{\delta\rho^h}{\mu\Xi}\frac{\partial\psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right) + \frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt} + \frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right] \\ & - \frac{p-c}{\Xi}\left\{\left[\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q}\right]\left(1+\frac{dp}{dt}\right) + \left[\pi\frac{\partial x_2^\ell}{\partial s} + (1-\pi)\frac{\partial x_2^h}{\partial s}\right]\frac{ds}{dt}\right\} \\ & - \frac{p_a - g_1'}{\Xi}\left(\frac{\partial a^h}{\partial p_a} + \frac{\partial a^\ell}{\partial p_a}\right)\frac{dp_a}{dt} \end{aligned}$$

where we have defined Ξ as

$$\Xi \equiv \left[\pi\frac{\partial x_2^\ell}{\partial q} + (1-\pi)\frac{\partial x_2^h}{\partial q}\right]\frac{dq}{dt} + \left[\pi\frac{\partial x_2^\ell}{\partial s} + (1-\pi)\frac{\partial x_2^h}{\partial s}\right]\frac{ds}{dt} + \pi\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt} + (1-\pi)\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt}$$

Proof. See Appendix C.4. ■

Once again, we get an expression for the optimal t which depends on three types of considerations. On the right hand side of the expression above, the terms appearing on the first line capture how self-selection considerations affect the optimal value for t. They depend both on the difference in consumption of good x_2 by a low-skilled and by a high-skilled mimicker, and on how a variation in t is going to affect, through the induced variations in a^h and s, the incentives for a high-skilled to behave as a mimicker. The terms contained in the second and third line capture the non-welfaristic motives affecting the optimal choice for t; these terms are due to the fact that the government launders the individual preferences when evaluating their utility. Finally, the terms contained in the last two lines capture the effect of a variation in t on the revenue obtained by the government through the tax levied on the profits of the monopolist selling the good x_2.

Understanding the sign of the expressions above is not simple. In fact, several terms are not well defined in sign.

Assuming that the net variation in the consumer price for good x_2 is positive, that is $1 + dp/dt > 0$, is not sufficient to state something else about the optimal tax rate formula.

The problem is twofold: we cannot define apriori the sign of term Ξ, and the net effect of t on the final consumed quantities x_2 for both agents, i.e. dx_2^i/dt. Hereafter we will discuss how to interpret the optimal formula when the platform answers by reducing data exploitation, $ds/dt < 0$, and by increasing price, $dp_a/dt > 0$, or by decreasing it, $dp_a/dt < 0$, but it is not enough to compensate the reduction in demand for ads, that is $da^i/dt < 0$.

Therefore, the first line assumes a negative expression, if the low-skilled one consumes more than high-skilled mimicker. In the previous section, this circumstance was obvious since persuasive advertising affected only low-skilled agents. Now, it is possible also the opposite case and self-selection considerations might end to call for a positive tax rate.

Furthermore, the second and third lines would favourite selecting a positive tax rate on the consumption of good x_2, because of “laundering" preferences by the government.

Lastly, the effect on fiscal revenue would be negative due to the hypothesis that the profits of the domestic firm selling good x_2 are fully taxed. This negative impact could be partially compensated (or enhanced) when the platform decreases (or increases) the price for ads.

4.4.1 Optimal advertising taxation with data exploitation

Suppose now that the government chooses to tax directly the ads that the monopolist buys instead of taxing the consumption of good x_2. Denote by t_a the specific tax on ads. The government's problem changes as follows:

$$\max_{I^\ell, B^\ell, I^h, B^h, t_a} \quad V^\ell\left(p, B^\ell, I^\ell; a, s\right) - \rho^\ell \psi\left(x_2^\ell, a, s\right)$$

subject to:

$$\begin{aligned}
&V^h\left(p, B^h, I^h; a, s\right) - \rho^h \psi\left(x_2^h, a, s\right) \geq \overline{V}, \\
&V^h\left(p, B^h, I^h; a, s\right) \geq V^h\left(p, B^\ell, I^\ell; a, s\right), \\
&\left(I^\ell - B^\ell\right)\pi + \left(I^h - B^h\right)(1-\pi) + \left(a^h + a^\ell\right)t_a + \\
&+\tau\left\{\left[\pi x_2^\ell\left(p, a^\ell, s\right) + (1-\pi)\, x_2^h\left(p, a^h, s\right)\right](p-c) - \left(a^h + a^\ell\right)(p_a + t_a)\right\} \geq \overline{R}.
\end{aligned}$$

Apart from the fact that the indirect utilities V^j (for $j = \ell, h$) no longer depend on t (since t is now by assumption equal to zero), the differences with the case considered in the previous subsection only pertain to the government's budget constraint. Here we have a specific tax t_a on ads bought by monopolist. The rate at which the tax is levied is assumed to be independent on whether the advertising is targeted towards high- or low-skilled agents.

While the maximization problem solved by the platform is still the one that we have considered in the previous subsection, the maximization problem solved by the monopolist selling good x_2 becomes the following:

$$\max_{p, a^h, a^\ell} \quad (1-\tau)\left\{\left[\pi x_2^\ell\left(p, a^\ell, s\right) + (1-\pi)\, x_2^h\left(p, a^h, s\right)\right](p-c) - \left(a^h + a^\ell\right)(p_a + t_a)\right\}$$

The first order conditions are:

$$\pi x_2^\ell + (1-\pi)\, x_2^h + \left[\pi \frac{\partial x_2^\ell}{\partial q} + (1-\pi)\,\frac{\partial x_2^h}{\partial q}\right](p-c) = 0 \tag{4.16}$$

$$(p-c)\,(1-\pi)\,\frac{\partial x_2^h}{\partial a^h} - (p_a + t_a) = 0 \tag{4.17}$$

$$(p-c)\,\pi \frac{\partial x_2^\ell}{\partial a^\ell} - (p_a + t_a) = 0 \tag{4.18}$$

Totally differentiating the three first order conditions above allows deriving expressions for

$$\begin{aligned}
\frac{dp}{dt_a} &= \frac{dp}{dt_a}\,|_{\substack{ds=0\\ dp_a=0}} + \frac{dp}{dp_a}\frac{dp_a}{dt_a} + \frac{dp}{ds}\frac{ds}{dt_a}\\
\frac{da^h}{dt_a} &= \frac{da^h}{dt_a}\,|_{\substack{ds=0\\ dp_a=0}} + \frac{da^h}{dp_a}\frac{dp_a}{dt_a} + \frac{da^h}{ds}\frac{ds}{dt_a}\\
\frac{da^\ell}{dt_a} &= \frac{da^\ell}{dt_a}\,|_{\substack{ds=0\\ dp_a=0}} + \frac{da^\ell}{dp_a}\frac{dp_a}{dt_a} + \frac{da^\ell}{ds}\frac{ds}{dt_a}
\end{aligned}$$

Notice that the first order conditions of the government's problem with respect to B^h and B^ℓ are still given by (4.14)-(4.15).

Proposition 12 *Given our assumption that utility is quasi-linear in x_1, at an optimum t_a will be set according to the following rule:*

$$\begin{aligned}
t_a = {} & \frac{\lambda}{\mu}\left(x_2^\ell - \widehat{x}_2\right)\frac{\frac{dp}{dt_a}}{\frac{da^\ell}{dt_a} + \frac{da^h}{dt_a}}\\
& + \frac{\rho^\ell \frac{\partial \psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}\left[\frac{\partial x_2^\ell}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt_a} + \frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt_a}\right] + \delta\rho^h \frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}\left[\frac{\partial x_2^h}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt_a} + \frac{\partial x_2^h}{\partial s}\frac{ds}{dt_a}\right]}{\mu\left[\frac{da^\ell}{dt_a} + \frac{da^h}{dt_a}\right]}\\
& - \frac{(p-c)\left\{\pi\left[\frac{\partial x_2^\ell}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt_a}\right] + (1-\pi)\left[\frac{\partial x_2^h}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^h}{\partial s}\frac{ds}{dt_a}\right]\right\}}{\frac{da^\ell}{dt_a} + \frac{da^h}{dt_a}}\\
& - \left(p_a - g_1'\right)\frac{\left(\frac{\partial a^h}{\partial p_a} + \frac{\partial a^\ell}{\partial p_a}\right)}{\left(\frac{da^h}{dt_a} + \frac{da^\ell}{dt_a}\right)}\frac{dp_a}{dt_a}
\end{aligned}$$

Proof. See Appendix C.5. ■

Hence, we get an expression for the optimal t_a which depends on three types of considerations. On the right hand side of the expression above, the terms appearing on the first line capture how self-selection considerations affect the optimal value for t_a. They depend both on the difference in consumption of good x_2 by a low-skilled and by a high-skilled mimicker, and on how a variation in t_a is going to affect, through the induced variations in a^h and s, the incentives for a high-skilled to behave as a mimicker. The terms contained in the second and third line capture the non-welfaristic motives affecting the optimal choice for t_a; these terms are due to the fact that the government launders the individual preferences when evaluating their utility. Finally, the terms contained in the last two lines capture the effect of a variation in t_a on the revenue obtained by the government through the tax levied on the profits of the monopolist selling the good x_2.

Leaving unchanged and adequately adapting the hypotheses made in the previous case, we discuss hereafter the above formula when $dp/dt_a > 0$, $da^i/dt_a < 0$, and $dx^i/dt_a < 0$.

Thus, the first line assumes a negative expression, if the low-skilled ones consume more than high-skilled mimickers. On the contrary, when mimickers consume more, self-selection considerations might end to call for a positive tax rate.

Furthermore, the second and third lines would favourite selecting a positive tax rate on the consumption of good x_2, because of "laundering" preferences by the government.

Lastly, the effect on fiscal revenue would be negative due to the hypothesis that the profits of the domestic firm selling good x_2 are fully taxed. This negative impact could be partially compensated (or enhanced) when the platform decreases (or increases) the price for ads.

4.4.2 Introducing t_a when t is optimally chosen

In this last paragraph, we deal with the issue of introducing a tax on data when commodity taxation has been already optimally chosen. Thus, we are at an equilibrium where t and $T(I)$ are optimally levied, but $t_a = 0$.
Suppose to have solved the government's problem:

$$\max_{I^\ell, B^\ell, I^h, B^h, t} \quad V^\ell\left(q, B^\ell, I^\ell; a^\ell, s\right) - \rho^\ell \psi\left(x_2^\ell, a^\ell, s\right)$$

subject to:

$$\begin{aligned} V^h\left(q, B^h, I^h; a^h, s\right) - \rho^h \psi\left(x_2^h, a^h, s\right) &\geq \overline{V}, \\ V^h\left(q, B^h, I^h; a^h, s\right) &\geq V^h\left(q, B^\ell, I^\ell; a^h, s\right), \end{aligned}$$

$$\begin{aligned} &\left(I^\ell - B^\ell + t x_2^\ell\right)\pi + \left(I^h - B^h + t x_2^h\right)(1-\pi) + t_a\left(a^h + a^\ell\right) \\ &+\tau\left\{\left[\pi x_2^\ell\left(q, a^\ell, s\right) + (1-\pi)\, x_2^h\left(q, a^h, s\right)\right](p-c) - \left(a^h + a^\ell\right)(p_a + t_a)\right\} \geq \overline{R}. \end{aligned}$$

Being at an initial equilibrium which is given by the solution to the above optimization problem, consider a tax reform that marginally raises t_a, starting from a value which is by assumption equal to zero, while at the same time varying t according to:

$$dt = -\frac{\frac{dp}{dt_a}}{1 + \frac{dp}{dt}} dt_a$$

where

$$\begin{aligned} \frac{dp}{dt_a} &\equiv \frac{dp}{dt_a}\Big|_{\substack{ds=0 \\ dp_a=0}} + \frac{dp}{dp_a}\frac{dp_a}{dt_a} + \frac{dp}{ds}\frac{ds}{dt_a} \\ \frac{dp}{dt} &\equiv \frac{dp}{dt}\Big|_{\substack{ds=0 \\ dp_a=0}} + \frac{dp}{dp_a}\frac{dp_a}{dt} + \frac{dp}{ds}\frac{ds}{dt} \end{aligned}$$

Notice that, by construction, the tax reform leaves unchanged the equilibrium value for $q = t + p(t, t_a; p_a, s)$. Also, the reform does not exert any effect on the self-selection constraint.

Proposition 13 *Defining* $\frac{dx_2^j}{dt_a}$ *as* $\frac{dx_2^j}{dt_a} \equiv \frac{\partial x_2^j}{\partial a^j}\frac{da^j}{dt_a} + \frac{\partial x_2^j}{\partial s}\frac{ds}{dt_a} + \frac{\partial x_2^j}{\partial q}\frac{dp}{dt_a}$ *and the elasticity* ϵ_{a^j,p_a} *as* $\epsilon_{a^j,p_a} \equiv \frac{da^j}{dp_a}\frac{p_a}{a^j}$*, we get the following expression, whose sign tells us whether introducing a small tax on advertising is welfare-enhancing or not:*

$$\left(q - c - \frac{\rho^\ell}{\mu\pi}\frac{\partial\psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}\right)\pi\frac{dx_2^\ell}{dt_a} + \left(q - c - \frac{\delta\rho^h}{\mu(1-\pi)}\frac{\partial\psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}\right)(1-\pi)\frac{dx_2^h}{dt_a}$$
$$- \sum_{j=\ell,h}\left(1 + \epsilon_{a^j,p_a}\right)a^j\frac{dp_a}{dt_a} + \frac{\lambda}{\mu}\left(\widehat{x}_2 - x_2^\ell\right)\frac{dp}{dt_a} \tag{4.19}$$

Proof. See Appendix C.7. ■

The sign of the expression (4.19) tells us whether, starting from an initial optimum where $\tau = 1$ and both t and $T(I)$ are optimally chosen, an introduction of a small tax on advertising is welfare-enhancing or not. In particular, a small tax on advertising will be welfare-enhancing if the sign of the above expression is positive; otherwise, it will be a small subsidy on advertising that will prove to be welfare-enhancing.

On the first line, the term $q - c - \frac{\rho^\ell}{\mu\pi}\frac{\partial\psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}$ can be interpreted as the difference between the consumer price for good 2 paid by low-skilled consumers and its "effective" marginal cost, which also includes the negative impact of a marginal increase in the consumption of good 2 when individual utilities are evaluated according to the laundered utility function used by the government. A similar interpretation applies for the term $q - c - \frac{\delta\rho^h}{(1-\pi)\mu}\frac{\partial\psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}$, relative to high-skilled consumers. The first term in the second line of the expression above captures instead the effect of the proposed tax reform on the revenue collected by the government through the taxation of the profits of the domestic firm selling the good x_2. Assuming that the tax t_a is partly shifted onto the platform, so that $dp_a/dt_a < 0$, the sign of this budget effect is positive (resp.: negative) if the elasticities ϵ_{a^j,p_a} (for $j = h, \ell$) are smaller (resp.: larger) than one in absolute value. The last term is a self-selection term that depends on the difference between the amount of good x_2 demanded by a true low-skilled and by a high-skilled mimicker. Assuming that $\rho^\ell > \rho^h$, i.e. that the behavior of low-skilled is more easily affected by persuasive advertising, and noticing that from the first order conditions (4.17)-(4.18) of the firm selling good x_2 we have that $\frac{\partial x_2^h}{\partial a^h} = \frac{\partial x_2^\ell}{\partial a^\ell}\pi/(1-\pi)$. One can conclude that, as long as the proportion of high-skilled agents is not too large (i.e. $\pi/(1-\pi)$ is not too small), a^h will be set at a value that is smaller than a^ℓ, and therefore $\widehat{x}_2 - x_2^\ell < 0$. Then, we will have that the sign of the self-selection term will be positive when $dp/dt_a < 0$.

4.5 Summary and concluding remarks

For what we know, the present work is the first attempt at formalizing, under the optimal taxation framework, the idea that persuasive advertising may distort individuals' utility in a way which is not reflected by government's preferences. We have tried to catch two aspects through an optimal taxation formula: the role of targeted advertising in changing the perception of utility derived by the consumption of the advertised good; and the distortion coming from monopoly.

By doing this, we have found that a positive commodity taxation may be justified, although our results hold under strict assumptions on preferences. The formulae that we have derived seek to correct three distortions: self-selection issues, persuasive advertising and market power.

In addition, we have considered the role of data in targeted advertising, evaluating how the non-welfarist terms can justify a positive tax on commodity or directly on ads.

Although it has not been possible to find a clear answer to the privacy issue since the effect of taxes on data exploitation is not well-defined, it seems plausible that it would decrease together with ads.

Lastly, we have shown how the implementation of a tax on advertising jointly with an optimal commodity tax might be welfare-enhancing, especially when mimicker is more sensitive to ads compared to low-skilled agents. It is an interesting and fascinating result if we think of targeting as a way to address customer better and in a more efficient way. Traditional ads may be more effective to low-skilled, whereas data exploitation may reverse the relative sensitivity of agents. Precisely why this hypothesis deserves empirical tests.

We are aware of the embedded limitations of the present analysis, but at the same time it can make governments reflect on opportunity to use fiscal tools to correct eventually undesired effects related to digital advertising and data exploitation, particularly when they care more than individuals about privacy protection for instance. Besides taxation, other instruments can be designed to address these issues, like new laws to limit data mining or the use of nudges, but it is out of the scope of present paper. We reserve to characterize jointly these aspects and optimal taxes in future works, as well as the capacity for firms to discriminate prices.

Said that, we hope that our exposition will serve to shed a new light on how to deal with the phenomenon of digitalization, and to widen its current research areas.

References

[1] Aronsson, T., and O. Johansson-Stenman (2018) *"Paternalism against Veblen: Optimal taxation and non-respected preferences for social comparisons"*, American Economic Journal: Economic Policy, 10 (1), 39-76.

[2] Atkinson, A. B., and J. E. Stiglitz (1976) *"The design of tax structure: direct versus indirect taxation"*, Journal of Public Economics, 6, 55-75.

[3] Becker, G. S., and K. M. Murphy (1993) *"A simple theory of advertising as a good or bad"* Quarterly Journal of Economics, 108 (4), 941-964.

[4] Blomquist, S., and L. Micheletto (2006) *"Optimal redistributive taxation when government's and agents' preferences differ"*, Journal of Public Economics, 90 (6-7), 1215-1233.

[5] del Mar Racionero, M. (2000) *"Optimal redistribution with unobservable preferences for an observable merit good"*, International Tax and Public Finance 7.4-5, 479-501.

[6] del Mar Racionero, M. (2001) *"Optimal Tax Mix with Merit Goods"*, Oxford University Press 53, 628-641.

[7] Farhi, E., and X. Gabaix (2020) *"Optimal taxation with behavioral agents"*, American Economic Review (forthcoming)

[8] Gerritsen, A. (2016) *"Optimal taxation when people do not maximize well-being"*, Journal of Public Economics, 144, 122-139.

[9] Kanbur, R., J. Pirttilä, and M. Tuomala (2006) *"Non-welfarist optimal taxation and behavioural public economics"*, Journal of Economic Surveys, 20(5), 849-868.

[10] Micheletto, L. (2011) *"Optimal nonlinear redistributive taxation and public good provision in an economy with Veblen effects"*, Journal of Public Economic Theory, 13 (1), 71-96.

[11] Mirrlees, J. A. (1971) *"An exploration in the theory of optimum income taxation"*, Review of Economic Studies, 38, 175-208.

[12] Mirrlees, J. A. (1976) *"Optimal tax theory: a synthesis"*, Journal of Public Economics, 6, 327-358.

[13] Saez, E. (2001) *"Using Elasticities to Derive Optimal Income Tax Rates"*, Review of Economic Studies, 68 (1), 205-229.

[14] Stiglitz, J. E. (1982) *"Self-selection and Pareto-efficient taxation"*, Journal of Public Economics, 17 (2), 213-240.

[15] Thaler, R. H., and C. R. Sunstein (2008) *"Nudge"*, Yale University Press.

Appendix A

Main derivations for Chapter 2

A.1 Derivation of comparative statics for a tax on platform's revenue

Starting from the FOCs of platform's revenue:

$$(1-t_r)xp_a - k_a = 0 \tag{A.1}$$

$$(1-t_r)xp_s - k_s = 0 \tag{A.2}$$

We can totally differentiate the above conditions and rewrite them into the matrix form:

$$\begin{bmatrix} (1-t_r)xp_{aa} - k_{aa} & (1-t_r)xp_{as} - k_{as} \\ (1-t_r)xp_{sa} - k_{sa} & (1-t_r)xp_{ss} - k_{ss} \end{bmatrix} \begin{bmatrix} da \\ ds \end{bmatrix} = \begin{bmatrix} xp_a dt_r \\ xp_s dt_r \end{bmatrix}$$

By assuming that the Hessian matrix is negative semidefinite. It can be expressed as: $\Omega_r = [(1-t_r)xp_{aa} - k_{aa}]\,[(1-t_r)xp_{ss} - k_{ss}] - [(1-t_r)xp_{as} - k_{as}]^2 > 0$.
Thus, it is possible to show the effect of this tax on platform's choices about a and s:

$$\Omega_r \frac{da}{dt_r} = xp_a[(1-t_r)xp_{ss} - k_{ss}] - xp_s[(1-t_r)xp_{as} - k_{as}] < 0 \tag{A.3}$$

$$\Omega_r \frac{ds}{dt_r} = xp_s[(1-t_r)xp_{aa} - k_{aa}] - xp_a[(1-t_r)xp_{as} - k_{as}] < 0 \tag{A.4}$$

A.2 Derivation of comparative statics for a VAT on platform's revenue

Starting from the FOCs of platform's revenue:

$$xp_a - k_a(1+\tau) = 0 \tag{A.5}$$

$$xp_s - k_s(1+\tau) = 0 \tag{A.6}$$

We can totally differentiate the above conditions and rewrite them into the matrix form:

$$\begin{bmatrix} xp_{aa} - k_{aa}(1+\tau) & xp_{as} - k_{as}(1+\tau) \\ \\ xp_{sa} - k_{sa}(1+\tau) & xp_{ss} - k_{ss}(1+\tau) \end{bmatrix} \begin{bmatrix} da \\ \\ ds \end{bmatrix} = \begin{bmatrix} k_a d\tau \\ \\ k_s d\tau \end{bmatrix}$$

By assuming that the Hessian matrix is negative semidefinite, i.e. $\Omega_\tau > 0$. Thus, it is possible to show the effect of this tax on platform's choices about a and s:

$$\Omega_\tau \frac{da}{dt_\tau} = k_a[xp_{ss} - k_{ss}(1+\tau)] - k_s[xp_{as} - k_{as}(1+\tau)] < 0 \tag{A.7}$$

$$\Omega_\tau \frac{ds}{dt_\tau} = k_s[xp_{aa} - k_{aa}(1+\tau)] - xp_a[xp_{as} - k_{as}(1+\tau)] < 0 \tag{A.8}$$

A.3 Derivation of comparative statics for a specific tax on ads

Starting from the FOCs of platform's revenue:

$$xp_a - k_a - t_a = 0 \tag{A.9}$$

$$xp_s - k_s = 0 \tag{A.10}$$

We can totally differentiate the above conditions and rewrite them into the matrix form:

$$\begin{bmatrix} xp_{aa} - k_{aa} & xp_{as} - k_{as} \\ \\ xp_{sa} - k_{sa} & xp_{ss} - k_{ss} \end{bmatrix} \begin{bmatrix} da \\ \\ ds \end{bmatrix} = \begin{bmatrix} dt_a \\ \\ 0 \end{bmatrix}$$

By assuming that the Hessian matrix is negative semidefinite, i.e. $\Omega_{t_a} > 0$. Thus, it is possible to show the effect of this tax on platform's choices about a and s:

$$\Omega_{t_a}\frac{da}{dt_a} = xp_{ss} - k_{ss} < 0 \tag{A.11}$$

$$\Omega_{t_a}\frac{ds}{dt_a} = -xp_{as} + k_{as} < 0 \tag{A.12}$$

A.4 Derivation of comparative statics for a specific tax on data

Starting from the FOCs of platform's revenue:

$$xp_a - k_a = 0 \tag{A.13}$$

$$xp_s - k_s - t_s = 0 \tag{A.14}$$

We can totally differentiate the above conditions and rewrite them into the matrix form:

$$\begin{bmatrix} xp_{aa} - k_{aa} & xp_{as} - k_{as} \\ xp_{sa} - k_{sa} & xp_{ss} - k_{ss} \end{bmatrix} \begin{bmatrix} da \\ ds \end{bmatrix} = \begin{bmatrix} 0 \\ dt_s \end{bmatrix}$$

By assuming that the Hessian matrix is negative semidefinite, i.e. $\Omega_{t_s} > 0$. Thus, it is possible to show the effect of this tax on platform's choices about a and s:

$$\Omega_{t_s}\frac{da}{dt_s} = -xp_{as} + k_{as} < 0 \tag{A.15}$$

$$\Omega_{t_s}\frac{ds}{dt_s} = xp_{aa} - k_{aa} < 0 \tag{A.16}$$

A.5 Derivation of comparative statics for a turnover tax in open economy

Starting from the FOCs of platform's revenue:

$$\begin{aligned}
(1-t_r)x^h p_a^h - k_a \frac{\partial a}{\partial a^h} &= 0 \\
(1-t_r)x^h p_s^h - k_s \frac{\partial s}{\partial s^h} &= 0 \\
x^f p_a^f - k_a \frac{\partial a}{\partial a^f} &= 0 \\
x^f p_s^f - k_s \frac{\partial s}{\partial a^f} &= 0
\end{aligned}$$

where $\frac{\partial s}{\partial s^i} = \frac{\partial s}{\partial s^i} = 1$ with $i = h, f$.

We can totally differentiate the above conditions and rewrite them into the matrix form:

$$\begin{bmatrix}
(1-t_r)x^h p_{aa} - k_{aa}\frac{\partial^2 a}{\partial a^{h2}} & -k_{aa}\frac{\partial^2 a}{\partial a^h \partial a^f} & (1-t_r)x^h p_{as} - k_{as}\frac{\partial a \partial s}{\partial a^h \partial s^h} & -k_{as}\frac{\partial a \partial s}{\partial a^h \partial s^f} \\
-k_{aa}\frac{\partial^2 a}{\partial a^{h2}} & x^f p_{aa} - k_{aa}\frac{\partial^2 a}{\partial a^h \partial a^f} & -k_{as}\frac{\partial a \partial s}{\partial a^f \partial s^h} & x^f p_{as} - k_{as}\frac{\partial a \partial s}{\partial a^f \partial s^f} \\
(1-t_r)x^h p_{as} - k_{as}\frac{\partial a \partial s}{\partial a^h \partial s^h} & -k_{as}\frac{\partial a \partial s}{\partial a^f \partial s^h} & (1-t_r)x^h p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{h2}} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \\
-k_{as}\frac{\partial a \partial s}{\partial a^h \partial s^f} & x^f p_{as} - k_{as}\frac{\partial a \partial s}{\partial a^f \partial s^f} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}}
\end{bmatrix}
* \begin{bmatrix} da^h \\ da^f \\ ds^h \\ ds^f \end{bmatrix} = \begin{bmatrix} x^h p_a^h dt_r \\ 0 \\ x^h p_s^h dt_r \\ 0 \end{bmatrix}$$

We assume that the Hessian matrix is negative semidefinite, i.e. $\Omega_{t_r}^{open} > 0$ and thereafter we consider only the case where marginal cost of ads k_a is constant and normalized to zero in order to avoid cumbersome calculations related to a matrix

4×4. Thus, it is possible to show the effects of this tax on platform's choices about a and s, home and abroad, result from the following:

$$\begin{bmatrix} (1-t_r)x^h p_{aa} & 0 & (1-t_r)x^h p_{as} & 0 \\ 0 & x^f p_{aa} & 0 & x^f p_{as} \\ (1-t_r)x^h p_{as} & 0 & (1-t_r)x^h p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{h2}} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \\ 0 & x^f p_{as} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix} \begin{bmatrix} da^h \\ da^f \\ ds^h \\ ds^f \end{bmatrix} = \begin{bmatrix} x^h p_a^h dt_r \\ 0 \\ x^h p_s^h dt_r \\ 0 \end{bmatrix}$$

Hence, one can derive immediately the effects of a tax on revenue on platform's choices at home and abroad:

$$
\begin{aligned}
\Omega_{t_r}^{open}\frac{da^h}{dt_r} &= x^h p_a^h * det\underbrace{\begin{bmatrix} x^f p_{aa} & 0 & x^f p_{as} \\ 0 & (1-t_r)x^h p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{h2}} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \\ x^f p_{as} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix}}_{(-)} \\
&+ x^h p_s^h * det\underbrace{\begin{bmatrix} 0 & (1-t_r)x^h p_{as} & 0 \\ x^f p_{aa} & 0 & x^f p_{as} \\ x^f p_{as} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix}}_{(+/-)} \gtrless 0 \\
\Omega_{t_r}^{open}\frac{da^f}{dt_r} &= -x^h p_a^h * det\underbrace{\begin{bmatrix} 0 & 0 & x^f p_{as} \\ (1-t_r)x^h p_{as} & (1-t_r)x^h p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{h2}} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \\ 0 & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix}}_{(+)} \\
&- x^h p_s^h * det\underbrace{\begin{bmatrix} (1-t_r)x^h p_{aa} & (1-t_r)x^h p_{as} & 0 \\ 0 & 0 & x^f p_{as} \\ 0 & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix}}_{(+)} < 0 \\
\Omega_{t_r}^{open}\frac{ds^h}{dt_r} &= x^h p_a^h * det\underbrace{\begin{bmatrix} 0 & x^f p_{aa} & x^f p_{as} \\ (1-t_r)x^h p_{as} & 0 & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \\ 0 & x^f p_{as} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix}}_{(+/-)} \\
&+ x^h p_s^h * det\underbrace{\begin{bmatrix} (1-t_r)x^h p_{aa} & 0 & 0 \\ 0 & x^f p_{aa} & x^f p_{as} \\ 0 & x^f p_{as} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix}}_{(-)} \gtrless 0 \\
\Omega_{t_r}^{open}\frac{ds^f}{dt_r} &= -x^h p_a^h * det\underbrace{\begin{bmatrix} 0 & x^f p_{aa} & 0 \\ (1-t_r)x^h p_{as} & 0 & (1-t_r)x^h p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{h2}} \\ 0 & x^f p_{as} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \end{bmatrix}}_{(+)} \\
&- x^h p_s^h * det\underbrace{\begin{bmatrix} (1-t_r)x^h p_{aa} & 0 & (1-t_r)x^h p_{as} \\ 0 & x^f p_{aa} & 0 \\ 0 & x^f p_{as} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \end{bmatrix}}_{(+)} < 0
\end{aligned}
$$

A.6 Derivation of comparative statics for a specific tax on ads or on data in open economy

By totally differentiating the first order conditions in presence of a tax on ads, t_a, and under the assumption that $k_a = 0$ and that the Hessian matrix is negative semidefinite, i.e. $\Omega_{t_a}^{open} = \Omega_{t_s}^{open} > 0$, we can rewrite them into the matrix form:

$$\begin{bmatrix} x^h p_{aa} & 0 & x^h p_{as} & 0 \\ 0 & x^f p_{aa} & 0 & x^f p_{as} \\ x^h p_{as} & 0 & x^h p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{h2}} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \\ 0 & x^f p_{as} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix} \begin{bmatrix} da^h \\ da^f \\ ds^h \\ ds^f \end{bmatrix} = \begin{bmatrix} dt_a \\ 0 \\ dt_s \\ 0 \end{bmatrix}$$

Hence, one can derive immediately the effects of a tax on advertising "intensity" on platform's choices at home and abroad:

$$\Omega_{t_a}^{open}\frac{da^h}{dt_a} = det\underbrace{\begin{bmatrix} x^f p_{aa} & 0 & x^f p_{as} \\ 0 & x^h p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{h2}} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \\ x^f p_{as} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix}}_{(-)} < 0$$

$$\Omega_{t_a}^{open}\frac{da^f}{dt_a} = -det\underbrace{\begin{bmatrix} 0 & 0 & x^f p_{as} \\ x^h p_{as} & x^h p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{h2}} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \\ 0 & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix}}_{(+)} < 0$$

$$\Omega_{t_a}^{open}\frac{ds^h}{dt_a} = det\underbrace{\begin{bmatrix} 0 & x^f p_{aa} & x^f p_{as} \\ x^h p_{as} & 0 & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \\ 0 & x^f p_{as} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix}}_{(+/-)} \gtrless 0$$

$$\Omega_{t_a}^{open}\frac{ds^f}{dt_a} = -det\underbrace{\begin{bmatrix} 0 & x^f p_{aa} & 0 \\ x^h p_{as} & 0 & x^h p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{h2}} \\ 0 & x^f p_{as} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \end{bmatrix}}_{(+)} < 0$$

Or, alternatively, one can derive immediately the effects of a tax on data exploitation on platform's choices at home and abroad:

$$\Omega_{t_s}^{open}\frac{da^h}{dt_s} = det\begin{bmatrix} 0 & x^h p_{as} & 0 \\ x^f p_{aa} & 0 & x^f p_{as} \\ x^f p_{as} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix} \gtrless 0$$
$$(+/-)$$

$$\Omega_{t_s}^{open}\frac{da^f}{dt_s} = -det\begin{bmatrix} x^h p_{aa} & x^h p_{as} & 0 \\ 0 & 0 & x^f p_{as} \\ 0 & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix} < 0$$
$$(+)$$

$$\Omega_{t_s}^{open}\frac{ds^h}{dt_s} = det\begin{bmatrix} x^h p_{aa} & 0 & 0 \\ 0 & x^f p_{aa} & x^f p_{as} \\ 0 & x^f p_{as} & x^f p_{ss} - k_{ss}\frac{\partial^2 s}{\partial s^{f2}} \end{bmatrix} < 0$$
$$(-)$$

$$\Omega_{t_s}^{open}\frac{ds^f}{dt_s} = -det\begin{bmatrix} x^h p_{aa} & 0 & x^h p_{as} \\ 0 & x^f p_{aa} & 0 \\ 0 & x^f p_{as} & -k_{ss}\frac{\partial^2 s}{\partial s^h \partial s^f} \end{bmatrix} < 0$$
$$(+)$$

Appendix B

Main derivations for Chapter 3

B.1 Comparative statics for a change in expectation about data exploitation

Totally differentiating the first order conditions of the platform, we get:

$$v_s ds - \psi_{s^e} - 2dZ - xp_a da - xp_s ds = 0$$

$$x(p_{aa}da + p_{as}ds)2n + xp_a\left[v_s ds - \psi_{s^e} ds^e + xp_a da + xp_s ds\right] - 2(k_{aa}da + k_{as}ds) = 0$$

$$(v_{ss}ds + xp_{as}da + xp_{ss}ds)2n + (v_s + xp_s)\left[v_s ds - \psi_{s^e} ds^e + p_a x da + xp_s ds\right] - 2(k_{as}da + k_{ss}ds) = 0$$

In matrix form:

$$\begin{bmatrix} -2 & -xp_a & v_s - xp_s \\ 0 & 2nxp_{aa} + x^2p_a^2 - 2k_{aa} & 2nxp_{as} + xp_a v_s + x^2 p_a p_s - 2k_{as} \\ 0 & 2nxp_{as} + (v_s + xp_s)xp_a - 2k_{as} & 2n(v_{ss} + xp_{ss}) + (v_s + xp_s)^2 - 2k_{ss} \end{bmatrix} \begin{bmatrix} dZ \\ da \\ ds \end{bmatrix} =$$

$$= \begin{bmatrix} \psi_{s^e} ds^e \\ xp_a \psi_{s^e} ds^e \\ (v_s + xp_s)\psi_{s^e} ds^e \end{bmatrix}$$

By assuming that conditions required for the Hessian matrix to be negative semidefinite hold, the determinant of the matrix above is defined as follows:

$$\Omega = -2 \left\{ \begin{array}{r} \left[2nxp_{aa} + x^2 p_a^2 - 2k_{aa}\right] \left[2n(v_{ss} + xp_{ss}) + (v_s + xp_s)^2 - 2k_{ss}\right] + \\ - \left[2nxp_{as} + (v_s + xp_s)xp_a - 2k_{as}\right]^2 \end{array} \right\} < 0$$

It is possible to show the effect of this change in users' expectation on platform's choices about Z, a and s:

$$\frac{dZ}{ds^e} = \frac{\left\{ \begin{array}{r} \psi_{s^e}(2nxp_{aa} + x^2 p_a^2 - 2k_{aa}) \left[2n(v_{ss} + xp_{ss}) + (v_s + xp_s)^2 - 2k_{ss}\right] + \\ -\psi_{s^e} \left[2nxp_{as} + xp_a v_s + x^2 p_a p_s - 2k_{as}\right]^2 + \\ +xp_a \psi_{s^e}(v_s - xp_s) \left[2nxp_{as} + (v_s + xp_s)xp_a - 2k_{as}\right] + \\ +x^2 p_a^2 \psi_{s^e} \left[2n(v_{ss} + xp_{ss}) + (v_s + xp_s)^2 - 2k_{ss}\right] + \\ -xp_a(v_s + xp_s)\psi_{s^e} + \left[2nxp_{as} + xp_a v_s + x^2 p_a p_s - 2k_{as}\right] + \\ -(v_s^2 - x^2 p_s^2)\psi_{s^e} \left[2nxp_{aa} + x^2 p_a^2 - 2k_{aa}\right] \end{array} \right\}}{\Omega} \gtrless 0$$

$$\frac{da}{ds^e} = \frac{\left\{ \begin{array}{r} -2xp_a \psi_{s^e} \left[2n(v_{ss} + xp_{ss}) + (v_s + xp_s)^2 - 2k_{ss}\right] + \\ +2(v_s + xp_s)\psi_{s^e} \left[2nxp_{as} + xp_a v_s + x^2 p_a p_s - 2k_{as}\right] \end{array} \right\}}{\Omega} < 0$$

$$\frac{ds}{ds^e} = \frac{\left\{ \begin{array}{r} -2(v_s + xp_s)\psi_{s^e} \left[2nxp_{aa} + x^2 p_a^2 - 2k_{aa}\right] + \\ +2xp_a \psi_{s^e} \left[2nxp_{as} + xp_a v_s + x^2 p_a p_s - 2k_{as}\right] \end{array} \right\}}{\Omega} < 0$$

B.2 Comparative statics after the introduction of a unit tax either on the access fee or on ads or on data

Totally differentiating the first order conditions of the platform, we get:

$$v_s ds - \psi_{s^e} - 2dZ - xp_a da - xp_s ds - dt_Z = 0$$

$$2nx(p_{aa}da + p_{as}ds) + xp_a\left[v_s ds - \psi_{s^e}ds^e - dZ\right] - k_{aa}da - k_{as}ds - dt_a = 0$$

$$2n(v_{ss}ds + xp_{as}da + xp_{ss}ds) + (v_s + xp_s)\left[v_s ds - \psi_{s^e}ds^e - dZ\right] - k_{as}da - k_{ss}ds - dt_s = 0$$

In matrix form:

$$\begin{bmatrix} -2 & -xp_a & v_s - xp_s \\ -xp_a & 2nxp_{aa} - k_{aa} & 2nxp_{as} + xp_a v_s - k_{as} \\ -(v_s + xp_s) & 2nxp_{as} - k_{as} & 2n(v_{ss} + xp_{ss}) + v_s(v_s + xp_s) - k_{ss} \end{bmatrix} \begin{bmatrix} dZ \\ da \\ ds \end{bmatrix} = $$
$$= \begin{bmatrix} \psi_{s^e}ds^e + dt_Z \\ xp_a\psi_{s^e}ds^e + dt_a \\ (v_s + xp_s)\psi_{s^e}ds^e + dt_s \end{bmatrix}$$

By assuming that conditions required for the Hessian matrix to be negative semidefinite hold, i.e. $\Omega < 0$, it is possible to show the effect following the introduction of a tax the access fee, t_Z, on platform's choices about Z, a and s:

$$\frac{dZ}{dt_Z} = \frac{\left\{ \begin{aligned} &(2nxp_{aa} - k_{aa})\left[2n(v_{ss} + xp_{ss}) + v_s(v_s + xp_s) - k_{ss}\right] + \\ &\qquad -(2nxp_{as} + xp_a v_s - k_{as})(2nxp_{as} - k_{as}) \end{aligned} \right\}}{\Omega} < 0$$

$$\frac{da}{dt_Z} = \frac{-(v_s + xp_s)(2nxp_{as} + xp_a v_s - k_{as}) + xp_a\left[2n(v_{ss} + xp_{ss}) + v_s(v_s + xp_s) - k_{ss}\right]}{\Omega} > 0$$

$$\frac{ds}{dt_Z} = \frac{-xp_a(2nxp_{as} - k_{as}) + (2nxp_{aa} - k_{aa})(v_s + xp_s)}{\Omega} > 0$$

or subsequent to a tax on ads, t_a:

$$\frac{dZ}{dt_a} = \frac{(v_s - xp_s)(2nxp_{as} - k_{as}) + xp_a\left[2n(v_{ss} + xp_{ss}) + v_s(v_s + xp_s) - k_{ss}\right]}{\Omega} \gtrless 0$$

$$\frac{da}{dt_a} = \frac{-2\left[2n(v_{ss} + xp_{ss}) + v_s(v_s + xp_s) - k_{ss}\right] + v_s^2 - x^2p_s^2}{\Omega} < 0$$

$$\frac{ds}{dt_a} = \frac{xp_a(v_s + xp_s) + 2(2nxp_{as} - k_{as})}{\Omega} < 0$$

and a tax on data, t_s:

$$\frac{dZ}{dt_s} = \frac{-xp_a(2nxp_{as} + xp_a v_s - k_{as}) - (v_s - xp_s)(2nxp_{aa} - k_{aa})}{\Omega} \gtrless 0$$

$$\frac{da}{dt_s} = \frac{-xp_a(v_s - xp_s) + 2(2nxp_{as} + xp_a v_s - k_{as})}{\Omega} \lesseqgtr 0$$

$$\frac{ds}{dt_s} = \frac{-2(2nxp_{aa} - k_{aa}) - x^2p_a^2}{\Omega} < 0$$

Appendix C

Main derivations for Chapter 4

C.1 Comparative statics for firm

Totally differentiating the set of first order conditions for the firm gives:

$$
\begin{aligned}
&\pi \frac{\partial x_2^\ell}{\partial B^\ell} dB^\ell + \pi \frac{\partial x_2^\ell}{\partial a} da + \pi \frac{\partial x_2^\ell}{\partial q} dp + \pi \frac{\partial x_2^\ell}{\partial q} dt + (1-\pi) \frac{\partial x_2^h}{\partial B^h} dB^h \\
&+ (1-\pi) \frac{\partial x_2^h}{\partial q} (dp + dt) + \left[\pi \frac{\partial x_2^\ell}{\partial q} + (1-\pi) \frac{\partial x_2^h}{\partial q} \right] dp + \\
&+ (p-c)\, \pi \frac{\partial^2 x_2^\ell}{\partial q \partial B^\ell} dB^\ell + (p-c)\, \pi \frac{\partial^2 x_2^\ell}{\partial q \partial a} da + (p-c)\, \pi \frac{\partial^2 x_2^\ell}{\partial q \partial q} (dp + dt) \\
&+ (p-c)(1-\pi) \frac{\partial^2 x_2^h}{\partial q \partial B^h} dB^h + (p-c)(1-\pi) \frac{\partial^2 x_2^h}{\partial q \partial q} (dp + dt) = 0
\end{aligned}
$$

$$
\pi \frac{\partial x_2^\ell}{\partial a} dp + (p-c)\, \pi \frac{\partial^2 x_2^\ell}{\partial a \partial B^\ell} dB^\ell + (p-c)\, \pi \frac{\partial^2 x_2^\ell}{\partial a \partial a} da + (p-c)\, \pi \frac{\partial^2 x_2^\ell}{\partial a \partial q} (dp + dt) - \theta''(a)\, da = 0
$$

In matrix form we have:

$$\begin{bmatrix} 2\left[\pi\frac{\partial x_2^\ell}{\partial q}+(1-\pi)\frac{\partial x_2^h}{\partial q}\right]+(p-c)\left[\pi\frac{\partial^2 x_2^\ell}{\partial q\partial q}+(1-\pi)\frac{\partial^2 x_2^h}{\partial q\partial q}\right] & \pi\left[\frac{\partial x_2^\ell}{\partial a}+(p-c)\frac{\partial^2 x_2^\ell}{\partial q\partial a}\right] \\ \pi\left[\frac{\partial x_2^\ell}{\partial a}+(p-c)\frac{\partial^2 x_2^\ell}{\partial a\partial q}\right] & (p-c)\,\pi\frac{\partial^2 x_2^\ell}{\partial a\partial a}-\theta''(a) \end{bmatrix}\begin{bmatrix} dp \\ da \end{bmatrix}$$

$$= \begin{bmatrix} \left\{ \begin{array}{c} -\pi\left[\frac{\partial x_2^\ell}{\partial B^\ell}+(p-c)\frac{\partial^2 x_2^\ell}{\partial q\partial B^\ell}\right]dB^\ell-(1-\pi)\left[\frac{\partial x_2^h}{\partial B^h}+(p-c)\frac{\partial^2 x_2^h}{\partial q\partial B^h}\right]dB^h \\ -\left[\left(\frac{\partial x_2^\ell}{\partial q}+(p-c)\frac{\partial^2 x_2^\ell}{\partial q\partial q}\right)\pi+\left((p-c)\frac{\partial^2 x_2^h}{\partial q\partial q}+\frac{\partial x_2^h}{\partial q}\right)(1-\pi)\right]dt \end{array} \right\} \\ -(p-c)\,\pi\left[\frac{\partial^2 x_2^\ell}{\partial a\partial B^\ell}dB^\ell+\frac{\partial^2 x_2^\ell}{\partial a\partial q}dt\right] \end{bmatrix}$$

Define the Hessian matrix Ω as

$$\begin{aligned} \Omega \equiv\ & \left\{\left[\pi\frac{\partial^2 x_2^\ell}{\partial q\partial q}+(1-\pi)\frac{\partial^2 x_2^h}{\partial q\partial q}\right](p-c)+\left(\pi\frac{\partial x_2^\ell}{\partial q}+(1-\pi)\frac{\partial x_2^h}{\partial q}\right)2\right\}\left[\pi\frac{\partial^2 x_2^\ell}{\partial a\partial a}(p-c)-\theta''(a)\right] \\ & -\pi^2\left[\frac{\partial x_2^\ell}{\partial a}+(p-c)\frac{\partial^2 x_2^\ell}{\partial q\partial a}\right]^2>0 \end{aligned}$$

We therefore get the following comparative statics results:

$$\begin{aligned} \Omega\frac{dp}{dt} =\ & \pi^2(p-c)\frac{\partial^2 x_2^\ell}{\partial q\partial a}\left[\frac{\partial x_2^\ell}{\partial a}+(p-c)\frac{\partial^2 x_2^\ell}{\partial q\partial a}\right] \\ & -\left[\pi\frac{\partial^2 x_2^\ell}{\partial a\partial a}(p-c)-\theta''(a)\right]\left[\pi\frac{\partial x_2^\ell}{\partial q}+(1-\pi)\frac{\partial x_2^h}{\partial q}\right] \\ & -\left[\pi\frac{\partial^2 x_2^\ell}{\partial a\partial a}(p-c)-\theta''(a)\right]\left[\pi\frac{\partial^2 x_2^\ell}{\partial q\partial q}+(1-\pi)\frac{\partial^2 x_2^h}{\partial q\partial q}\right](p-c) \end{aligned}$$

$$\begin{aligned} \Omega\frac{da}{dt} =\ & \left[\pi\frac{\partial x_2^\ell}{\partial q}+(1-\pi)\frac{\partial x_2^h}{\partial q}\right]\left[\frac{\partial x_2^\ell}{\partial a}-(p-c)\frac{\partial^2 x_2^\ell}{\partial q\partial a}\right]\pi \\ & +(p-c)\frac{\partial x_2^\ell}{\partial a}\left[\pi\frac{\partial^2 x_2^\ell}{\partial q\partial q}+(1-\pi)\frac{\partial^2 x_2^h}{\partial q\partial q}\right]\pi<0 \end{aligned}$$

$$\Omega\frac{dp}{dB^h}=-(1-\pi)\left[\frac{\partial x_2^h}{\partial B^h}+(p-c)\frac{\partial^2 x_2^h}{\partial B^h\partial q}\right]\left[\pi\frac{\partial^2 x_2^\ell}{\partial a\partial a}(p-c)-\theta''(a)\right]\geq 0$$

$$\Omega\frac{da}{dB^h}=\pi(1-\pi)\left[\frac{\partial x_2^h}{\partial B^h}+(p-c)\frac{\partial^2 x_2^h}{\partial B^h\partial q}\right]\left[\frac{\partial x_2^\ell}{\partial a}+(p-c)\frac{\partial^2 x_2^\ell}{\partial q\partial a}\right]\geq 0$$

$$\Omega\frac{dp}{dB^{\ell}} = \pi^2\left(p-c\right)\frac{\partial^2 x_2^{\ell}}{\partial B^{\ell}\partial a}\left[\frac{\partial x_2^{\ell}}{\partial a}+\left(p-c\right)\frac{\partial^2 x_2^{\ell}}{\partial q\partial a}\right] -\pi\left[\frac{\partial x_2^{\ell}}{\partial B^{\ell}}+\left(p-c\right)\frac{\partial^2 x_2^{\ell}}{\partial B^{\ell}\partial q}\right]\left[\pi\frac{\partial^2 x_2^{\ell}}{\partial a\partial a}\left(p-c\right)-\theta''\left(a\right)\right]\geq 0$$

$$\Omega\frac{da}{dB^{\ell}} = \pi^2\left[\frac{\partial x_2^{\ell}}{\partial a}+\left(p-c\right)\frac{\partial^2 x_2^{\ell}}{\partial q\partial a}\right]\left[\frac{\partial x_2^{\ell}}{\partial B^{\ell}}+\left(p-c\right)\frac{\partial^2 x_2^{\ell}}{\partial B^{\ell}\partial q}\right] -2\pi\left(p-c\right)\frac{\partial^2 x_2^{\ell}}{\partial B^{\ell}\partial a}\left[\pi\frac{\partial x_2^{\ell}}{\partial q}+\left(1-\pi\right)\frac{\partial x_2^{h}}{\partial q}\right] -\pi\left(p-c\right)^2\frac{\partial^2 x_2^{\ell}}{\partial B^{\ell}\partial a}\left[\pi\frac{\partial^2 x_2^{\ell}}{\partial q\partial q}+\left(1-\pi\right)\frac{\partial^2 x_2^{h}}{\partial q\partial q}\right]\geq 0$$

C.2 Derivation of the optimal commodity tax rate

Denote the Lagrange multipliers associated to the three constraints above by, respectively, δ, λ and μ. Denote by a "hat" a variable pertaining to a high-skilled behaving as a mimicker. Define Δ as:

$$\Delta \equiv \frac{\partial V^{\ell}}{\partial q}-\frac{\partial \psi}{\partial x_2^{\ell}}\frac{\partial x_2^{\ell}}{\partial q}+\left(\delta+\lambda\right)\frac{\partial V^{h}}{\partial q}-\lambda\frac{\partial \widehat{V}}{\partial q}+\mu t\left[\pi\frac{\partial x_2^{\ell}}{\partial q}+\left(1-\pi\right)\frac{\partial x_2^{h}}{\partial q}\right] \tag{C.1}$$

First order condition with respect to B^h:

$$\left(\delta+\lambda\right)\frac{\partial V^{h}}{\partial B^{h}}+\mu\left[t\frac{\partial x_2^{h}}{\partial B^{h}}-1\right]\left(1-\pi\right)+\left[\mu t\pi-\frac{\partial \psi}{\partial x_2^{\ell}}\right]\frac{\partial x_2^{\ell}}{\partial a}\frac{da}{dB^{h}}+\Delta\frac{dp}{dB^{h}}=0$$

First order condition with respect to B^{ℓ}:

$$\frac{\partial V^{\ell}}{\partial B^{\ell}}-\frac{\partial \psi}{\partial x_2^{\ell}}\frac{\partial x_2^{\ell}}{\partial B^{\ell}}-\lambda\frac{\partial \widehat{V}}{\partial B^{\ell}}+\mu\left[t\frac{\partial x_2^{\ell}}{\partial B^{\ell}}-1\right]\pi+\left[\mu t\pi-\frac{\partial \psi}{\partial x_2^{\ell}}\right]\frac{\partial x_2^{\ell}}{\partial a}\frac{da}{dB^{\ell}}+\Delta\frac{dp}{dB^{\ell}}=0$$

First order condition with respect to t:

$$\frac{\partial V^{\ell}}{\partial q}-\frac{\partial \psi}{\partial x_2^{\ell}}\frac{\partial x_2^{\ell}}{\partial q}+(\delta+\lambda)\frac{\partial V^{h}}{\partial q}-\lambda\frac{\partial \widehat{V}}{\partial q}+\mu t\left[\pi\frac{\partial x_2^{\ell}}{\partial q}+(1-\pi)\frac{\partial x_2^{h}}{\partial q}\right]$$
$$+\mu\left[\pi x_2^{\ell}+(1-\pi)x_2^{h}\right]+\left[\mu t\pi-\frac{\partial \psi}{\partial x_2^{\ell}}\right]\frac{\partial x_2^{\ell}}{\partial a}\frac{da}{dt}+\Delta\frac{dp}{dt}=0 \qquad \text{(C.2)}$$

Denote Hicksian (compensated) demands by a "tilde" symbol; applying Roy's identity[1] and Slutsky equation[2], we can rewrite (C.2) as:

$$-\frac{\partial \psi}{\partial x_2^{\ell}}\frac{\partial \widetilde{x}_2^{\ell}}{\partial q}+\lambda\widehat{x}_2\frac{\partial \widehat{V}}{\partial B^{\ell}}+\mu t\left[\pi\frac{\partial \widetilde{x}_2^{\ell}}{\partial q}+(1-\pi)\frac{\partial \widetilde{x}_2^{h}}{\partial q}\right]+\left[\mu t\pi-\frac{\partial \psi}{\partial x_2^{\ell}}\right]\frac{\partial x_2^{\ell}}{\partial a}\frac{da}{dt}+\Delta\frac{dp}{dt}$$
$$= x_2^{\ell}\frac{\partial V^{\ell}}{\partial B^{\ell}}-\frac{\partial \psi}{\partial x_2^{\ell}}\frac{\partial x_2^{\ell}}{\partial B^{\ell}}x_2^{\ell}+(\delta+\lambda)x_2^{h}\frac{\partial V^{h}}{\partial B^{h}}+\mu t\left[\pi x_2^{\ell}\frac{\partial x_2^{\ell}}{\partial B^{\ell}}+(1-\pi)x_2^{h}\frac{\partial x_2^{h}}{\partial B^{h}}\right]$$
$$-\mu\left[\pi x_2^{\ell}+(1-\pi)x_2^{h}\right] \qquad \text{(C.3)}$$

Rewrite the first order conditions with respect to B^h and B^{ℓ} by multiplying for x_2^h and x_2^{ℓ} as, respectively:

$$(\delta+\lambda)\frac{\partial V^{h}}{\partial B^{h}}x_2^{h}+\mu\left[tx_2^{h}\frac{\partial x_2^{h}}{\partial B^{h}}-x_2^{h}\right](1-\pi)=-\left[\mu t\pi-\frac{\partial \psi}{\partial x_2^{\ell}}\right]\frac{\partial x_2^{\ell}}{\partial a}\frac{da}{dB^{h}}x_2^{h}-\Delta\frac{dp}{dB^{h}}x_2^{h} \qquad \text{(C.4)}$$

$$\frac{\partial V^{\ell}}{\partial B^{\ell}}x_2^{\ell}-\frac{\partial \psi}{\partial x_2^{\ell}}\frac{\partial x_2^{\ell}}{\partial B^{\ell}}x_2^{\ell}+\mu\left[tx_2^{\ell}\frac{\partial x_2^{\ell}}{\partial B^{\ell}}-x_2^{\ell}\right]\pi=\lambda\frac{\partial \widehat{V}}{\partial B^{\ell}}x_2^{\ell}-\left[\mu t\pi-\frac{\partial \psi}{\partial x_2^{\ell}}\right]\frac{\partial x_2^{\ell}}{\partial a}\frac{da}{dB^{\ell}}x_2^{\ell}-\Delta\frac{dp}{dB^{\ell}}x_2^{\ell} \qquad \text{(C.5)}$$

[1] It is $x_2(q,B)=-\frac{\partial V/\partial q}{\partial V/\partial B}$.

[2] It is $\frac{\partial x_2}{\partial q}=\frac{\partial \tilde{x}_2}{\partial q}-x_2\frac{\partial x_2}{\partial B}$.

Then we can use (C.4)-(C.5) to rewrite (C.3) as:

$$
\begin{aligned}
&-\frac{\partial \psi}{\partial x_2^\ell}\frac{\partial \widetilde{x}_2^\ell}{\partial q} + \lambda \widehat{x}_2 \frac{\partial \widehat{V}}{\partial B^\ell} + \mu t \left[\pi \frac{\partial \widetilde{x}_2^\ell}{\partial q} + (1-\pi)\frac{\partial \widetilde{x}_2^h}{\partial q}\right] + \left[\mu t \pi - \frac{\partial \psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dt} + \Delta \frac{dp}{dt} \\
= \; & \lambda \frac{\partial \widehat{V}}{\partial B^\ell} x_2^\ell - \left[\mu t \pi - \frac{\partial \psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dB^\ell} x_2^\ell - \Delta \frac{dp}{dB^\ell} x_2^\ell - \left[\mu t \pi - \frac{\partial \psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dB^h} x_2^h - \Delta \frac{dp}{dB^h} x_2^h,
\end{aligned} \tag{C.6}
$$

or, equivalently:

$$
\begin{aligned}
& \Delta \left[\frac{dp}{dt} + \frac{dp}{dB^\ell} x_2^\ell + \frac{dp}{dB^h} x_2^h\right] \\
= \; & \\
& \lambda \frac{\partial \widehat{V}}{\partial B^\ell}\left(x_2^\ell - \widehat{x}_2\right) + \frac{\partial \psi}{\partial x_2^\ell}\left\{\frac{\partial \widetilde{x}_2^\ell}{\partial q} + \left[\frac{da}{dt} + \frac{da}{dB^\ell} x_2^\ell + \frac{da}{dB^h} x_2^h\right]\frac{\partial x_2^\ell}{\partial a}\right\} \\
& -\mu t \left\{\left[\pi \frac{\partial \widetilde{x}_2^\ell}{\partial q} + (1-\pi)\frac{\partial \widetilde{x}_2^h}{\partial q}\right] + \pi \left[\frac{da}{dt} + \frac{da}{dB^\ell} x_2^\ell + \frac{da}{dB^h} x_2^h\right]\frac{\partial x_2^\ell}{\partial a}\right\}
\end{aligned} \tag{C.7}
$$

Now rewrite Δ as defined in (C.1) as:

$$
\begin{aligned}
\Delta \;\equiv\; & -\frac{\partial V^\ell}{\partial B^\ell} x_2^\ell + \frac{\partial \psi}{\partial x_2^\ell}\frac{\partial x_2^\ell}{\partial B^\ell} x_2^\ell - \mu t \pi x_2^\ell \frac{\partial x_2^\ell}{\partial B^\ell} \\
& - (\delta + \lambda)\frac{\partial V^h}{\partial B^h} x_2^h - \mu t (1-\pi) x_2^h \frac{\partial x_2^h}{\partial B^h} \\
& + \mu t \left[(1-\pi)\frac{\partial \widetilde{x}_2^h}{\partial q} + \pi \frac{\partial \widetilde{x}_2^\ell}{\partial q}\right] + \lambda \frac{\partial \widehat{V}}{\partial B^\ell}\widehat{x}_2 - \frac{\partial \psi}{\partial x_2^\ell}\frac{\partial \widetilde{x}_2^\ell}{\partial q}
\end{aligned}
$$

Using (C.4)-(C.5) we can re-express Δ as:

$$
\begin{aligned}
\Delta \;\equiv\; & -\mu \pi x_2^\ell - \lambda \frac{\partial \widehat{V}}{\partial B^\ell} x_2^\ell + \Delta \frac{dp}{dB^\ell} x_2^\ell + \left[\mu t \pi - \frac{\partial \psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dB^\ell} x_2^\ell \\
& -\mu (1-\pi) x_2^h + \Delta \frac{dp}{dB^h} x_2^h + \left[\mu t \pi - \frac{\partial \psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\frac{da}{dB^h} x_2^h \\
& + \mu t \left[(1-\pi)\frac{\partial \widetilde{x}_2^h}{\partial q} + \pi \frac{\partial \widetilde{x}_2^\ell}{\partial q}\right] + \lambda \frac{\partial \widehat{V}}{\partial B^\ell}\widehat{x}_2 - \frac{\partial \psi}{\partial x_2^\ell}\frac{\partial \widetilde{x}_2^\ell}{\partial q}
\end{aligned}
$$

Therefore, we can write:

$$\begin{aligned}\Delta\left[1-\frac{dp}{dB^\ell}x_2^\ell-\frac{dp}{dB^h}x_2^h\right] &= \mu t\left[(1-\pi)\frac{\partial\widetilde{x}_2^h}{\partial q}+\pi\frac{\partial\widetilde{x}_2^\ell}{\partial q}\right]+\lambda\frac{\partial\widehat{V}}{\partial B^\ell}\left(\widehat{x}_2-x_2^\ell\right)\\ &\quad -\frac{\partial\psi}{\partial x_2^\ell}\frac{\partial\widetilde{x}_2^\ell}{\partial q}-\mu\left[\pi x_2^\ell+(1-\pi)x_2^h\right]\\ &\quad +\left[\mu t\pi-\frac{\partial\psi}{\partial x_2^\ell}\right]\left[\frac{da}{dB^\ell}x_2^\ell+\frac{da}{dB^h}x_2^h\right]\frac{\partial x_2^\ell}{\partial a},\end{aligned}$$

or, equivalently:

$$\begin{aligned}\Delta &= \left[1-\frac{dp}{dB^\ell}x_2^\ell-\frac{dp}{dB^h}x_2^h\right]^{-1}\left\{\mu t\left[(1-\pi)\frac{\partial\widetilde{x}_2^h}{\partial q}+\pi\frac{\partial\widetilde{x}_2^\ell}{\partial q}\right]+\lambda\frac{\partial\widehat{V}}{\partial B^\ell}\left(\widehat{x}_2-x_2^\ell\right)-\frac{\partial\psi}{\partial x_2^\ell}\frac{\partial\widetilde{x}_2^\ell}{\partial q}\right\}\\ &\quad -\left[1-\frac{dp}{dB^\ell}x_2^\ell-\frac{dp}{dB^h}x_2^h\right]^{-1}\mu\left[\pi x_2^\ell+(1-\pi)x_2^h\right]\\ &\quad +\left[1-\frac{dp}{dB^\ell}x_2^\ell-\frac{dp}{dB^h}x_2^h\right]^{-1}\left[\frac{da}{dB^\ell}x_2^\ell+\frac{da}{dB^h}x_2^h\right]\left[\mu t\pi-\frac{\partial\psi}{\partial x_2^\ell}\right]\frac{\partial x_2^\ell}{\partial a}\end{aligned}$$

Substituting the expression above into (C.7) and collecting terms gives:

$$\begin{aligned}&-\frac{\frac{dp}{dt}+\frac{dp}{dB^\ell}x_2^\ell+\frac{dp}{dB^h}x_2^h}{1-\frac{dp}{dB^\ell}x_2^\ell-\frac{dp}{dB^h}x_2^h}\mu\left[\pi x_2^\ell+(1-\pi)x_2^h\right]\\ =&\\ &\frac{\lambda\frac{\partial\widehat{V}}{\partial B^\ell}\left(x_2^\ell-\widehat{x}_2\right)\left(1+\frac{dp}{dt}\right)}{1-\frac{dp}{dB^\ell}x_2^\ell-\frac{dp}{dB^h}x_2^h}-\frac{\mu t\left[\pi\frac{\partial\widetilde{x}_2^\ell}{\partial q}+(1-\pi)\frac{\partial\widetilde{x}_2^h}{\partial q}\right]\left(1+\frac{dp}{dt}\right)}{1-\frac{dp}{dB^\ell}x_2^\ell-\frac{dp}{dB^h}x_2^h}\\ &+\frac{\frac{\partial\psi}{\partial x_2^\ell}\frac{\partial\widetilde{x}_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)}{1-\frac{dp}{dB^\ell}x_2^\ell-\frac{dp}{dB^h}x_2^h}\\ &-\left[\mu t\pi-\frac{\partial\psi}{\partial x_2^\ell}\right]\left[\frac{da}{dt}+\frac{\left(1+\frac{dp}{dt}\right)\frac{da}{dB^\ell}x_2^\ell}{1-\frac{dp}{dB^\ell}x_2^\ell-\frac{dp}{dB^h}x_2^h}+\frac{\left(1+\frac{dp}{dt}\right)\frac{da}{dB^h}x_2^h}{1-\frac{dp}{dB^\ell}x_2^\ell-\frac{dp}{dB^h}x_2^h}\right]\frac{\partial x_2^\ell}{\partial a},\end{aligned}$$

or, equivalently:

$$
\begin{aligned}
& \mu t \left[\pi \frac{\partial \widetilde{x}_2^\ell}{\partial q} + (1-\pi) \frac{\partial \widetilde{x}_2^h}{\partial q} \right] \left(1 + \frac{dp}{dt} \right) \\
= \; & \lambda \frac{\partial \widehat{V}}{\partial B^\ell} \left(x_2^\ell - \widehat{x}_2 \right) \left(1 + \frac{dp}{dt} \right) \\
& + \frac{\partial \psi}{\partial x_2^\ell} \frac{\partial \widetilde{x}_2^\ell}{\partial q} \left(1 + \frac{dp}{dt} \right) + \mu \left[\pi x_2^\ell + (1-\pi) x_2^h \right] \left[\frac{dp}{dt} + \frac{dp}{dB^\ell} x_2^\ell + \frac{dp}{dB^h} x_2^h \right] \\
& - \left[\frac{da}{dB^\ell} x_2^\ell + \frac{da}{dB^h} x_2^h \right] \left(1 + \frac{dp}{dt} \right) \left[\mu t \pi - \frac{\partial \psi}{\partial x_2^\ell} \right] \frac{\partial x_2^\ell}{\partial a} \\
& - \left[\mu t \pi - \frac{\partial \psi}{\partial x_2^\ell} \right] \frac{\partial x_2^\ell}{\partial a} \frac{da}{dt} \left[1 - \frac{dp}{dB^\ell} x_2^\ell - \frac{dp}{dB^h} x_2^h \right]. \qquad \text{(C.8)}
\end{aligned}
$$

Noticing that from (4.5) we have

$$
\pi x_2^\ell + (1-\pi) x_2^h = - \left[\pi \frac{\partial x_2^\ell}{\partial q} + (1-\pi) \frac{\partial x_2^h}{\partial q} \right] (p - c),
$$

we can rewrite (C.8) as

$$
\begin{aligned}
& \mu t \left[\pi \frac{\partial \widetilde{x}_2^\ell}{\partial q} + (1-\pi) \frac{\partial \widetilde{x}_2^h}{\partial q} \right] \left(1 + \frac{dp}{dt} \right) \\
& + \mu t \left[\left(\frac{da}{dB^\ell} x_2^\ell + \frac{da}{dB^h} x_2^h \right) \left(1 + \frac{dp}{dt} \right) + \frac{da}{dt} \left(1 - \frac{dp}{dB^\ell} x_2^\ell - \frac{dp}{dB^h} x_2^h \right) \right] \pi \frac{\partial x_2^\ell}{\partial a} \\
= \; & \lambda \frac{\partial \widehat{V}}{\partial B^\ell} \left(x_2^\ell - \widehat{x}_2 \right) \left(1 + \frac{dp}{dt} \right) \\
& + \frac{\partial \psi}{\partial x_2^\ell} \frac{\partial \widetilde{x}_2^\ell}{\partial q} \left(1 + \frac{dp}{dt} \right) - \mu (p - c) \left[\pi \frac{\partial x_2^\ell}{\partial q} + (1-\pi) \frac{\partial x_2^h}{\partial q} \right] \left[\frac{dp}{dt} + \frac{dp}{dB^\ell} x_2^\ell + \frac{dp}{dB^h} x_2^h \right] \\
& + \left\{ \left[\frac{da}{dB^\ell} x_2^\ell + \frac{da}{dB^h} x_2^h \right] \left(1 + \frac{dp}{dt} \right) + \frac{da}{dt} \left[1 - \frac{dp}{dB^\ell} x_2^\ell - \frac{dp}{dB^h} x_2^h \right] \right\} \frac{\partial \psi}{\partial x_2^\ell} \frac{\partial x_2^\ell}{\partial a}.
\end{aligned}
$$

To simplify the interpretation of the optimality condition that we have derived above, assume that the agents' utility function is quasi-linear in x_1 so that $\partial x_2^j / \partial q = \partial \widetilde{x}_2^j / \partial q$ for $j = \ell, h$ and $da/dB^\ell = da/dB^h = dp/dB^\ell = dp/dB^h = 0$. Then, the condition above simplifies to:

$$\lambda \frac{\partial \widehat{V}}{\partial B^{\ell}} \left(\widehat{x}_2 - x_2^{\ell}\right) \left(1 + \frac{dp}{dt}\right) dt$$
$$+ \frac{\partial \psi}{\partial x_2^{\ell}} \left[\frac{\partial x_2^{\ell}}{\partial a} \frac{da}{dt} + \frac{\partial x_2^{\ell}}{\partial q} \left(1 + \frac{dp}{dt}\right)\right] dt$$
$$- \mu \left[\pi \frac{\partial x_2^{\ell}}{\partial q} + (1 - \pi) \frac{\partial x_2^{h}}{\partial q}\right] (p - c) \frac{dp}{dt} dt$$

Using the simplified expressions above, we get:

$$\begin{aligned} & \mu t \left\{ \left[\pi \frac{\partial \widetilde{x}_2^{\ell}}{\partial q} + (1 - \pi) \frac{\partial \widetilde{x}_2^{h}}{\partial q}\right] \left(1 + \frac{dp}{dt}\right) + \pi \frac{\partial x_2^{\ell}}{\partial a} \frac{da}{dt} \right\} \\ = & \\ & \lambda \frac{\partial \widehat{V}}{\partial B^{\ell}} \left(x_2^{\ell} - \widehat{x}_2\right) \left(1 + \frac{dp}{dt}\right) \\ & + \left[\frac{\partial x_2^{\ell}}{\partial a} \frac{da}{dt} + \frac{\partial \widetilde{x}_2^{\ell}}{\partial q} \left(1 + \frac{dp}{dt}\right)\right] \frac{\partial \psi}{\partial x_2^{\ell}} \\ & - \mu \left[\pi \frac{\partial \widetilde{x}_2^{\ell}}{\partial q} + (1 - \pi) \frac{\partial \widetilde{x}_2^{h}}{\partial q}\right] (p - c) \frac{dp}{dt}. \end{aligned} \tag{C.9}$$

where

$$1 + \frac{dp}{dt} = \frac{\left(\pi \frac{\partial x_2^{\ell}}{\partial q} + (1 - \pi) \frac{\partial x_2^{h}}{\partial q}\right) \left[\pi \frac{\partial^2 x_2^{\ell}}{\partial a \partial a} (p - c) - \theta''(a)\right] - \pi^2 \left[\frac{\partial x_2^{\ell}}{\partial a} + \frac{\partial^2 x_2^{\ell}}{\partial q \partial a} (p - c)\right] \frac{\partial x_2^{\ell}}{\partial a}}{\Omega}.$$

Finally, defining Y as

$$\mathrm{Y} \equiv \left[\pi \frac{\partial \widetilde{x}_2^{\ell}}{\partial q} + (1 - \pi) \frac{\partial \widetilde{x}_2^{h}}{\partial q}\right] \left(1 + \frac{dp}{dt}\right) + \pi \frac{\partial x_2^{\ell}}{\partial a} \frac{da}{dt},$$

we can finally restate (C.9) as

$$\begin{aligned} t &= \frac{\lambda}{\mu Y} \frac{\partial \widehat{V}}{\partial B^{\ell}} \left(x_2^{\ell} - \widehat{x}_2 \right) \left(1 + \frac{dp}{dt} \right) \\ &+ \frac{1}{\mu Y} \left[\frac{\partial x_2^{\ell}}{\partial a} \frac{da}{dt} + \frac{\partial \widetilde{x}_2^{\ell}}{\partial q} \left(1 + \frac{dp}{dt} \right) \right] \frac{\partial \psi}{\partial x_2^{\ell}} \\ &- \frac{p-c}{Y} \left[\pi \frac{\partial \widetilde{x}_2^{\ell}}{\partial q} + (1-\pi) \frac{\partial \widetilde{x}_2^{h}}{\partial q} \right] \frac{dp}{dt}. \end{aligned} \tag{C.10}$$

C.3 Comparative statics for platform and monopolist with a tax on commodity

Here we show the expression needed to understand the effect of a tax on commodity through comparative statics. Totally differentiating the set of first order condition of monopolist gives:

$$\begin{aligned} &\pi \frac{\partial x_2^{\ell}}{\partial B^{\ell}} dB^{\ell} + \pi \frac{\partial x_2^{\ell}}{\partial a^{\ell}} da^{\ell} + \pi \frac{\partial x_2^{\ell}}{\partial q} dp + \pi \frac{\partial x_2^{\ell}}{\partial q} dt + \pi \frac{\partial x_2^{\ell}}{\partial s} ds + (1-\pi) \frac{\partial x_2^{h}}{\partial B^{h}} dB^{h} \\ &+ (1-\pi) \frac{\partial x_2^{h}}{\partial a^{h}} da^{h} + (1-\pi) \frac{\partial x_2^{h}}{\partial q} (dp + dt) + (1-\pi) \frac{\partial x_2^{h}}{\partial s} ds \\ &+ \left[\pi \frac{\partial x_2^{\ell}}{\partial q} + (1-\pi) \frac{\partial x_2^{h}}{\partial q} \right] dp + (p-c) \pi \frac{\partial^2 x_2^{\ell}}{\partial q \partial B^{\ell}} dB^{\ell} \\ &+ (p-c) \pi \frac{\partial^2 x_2^{\ell}}{\partial q \partial a^{\ell}} da^{\ell} + (p-c) \pi \frac{\partial^2 x_2^{\ell}}{\partial q \partial s} ds + (p-c) \pi \frac{\partial^2 x_2^{\ell}}{\partial q \partial q} (dp + dt) \\ &+ (p-c)(1-\pi) \frac{\partial^2 x_2^{h}}{\partial q \partial B^{h}} dB^{h} + (p-c)(1-\pi) \frac{\partial^2 x_2^{h}}{\partial q \partial a^{h}} da^{h} + (p-c)(1-\pi) \frac{\partial^2 x_2^{h}}{\partial q \partial s} ds \\ &+ (p-c)(1-\pi) \frac{\partial^2 x_2^{h}}{\partial q \partial q} (dp + dt) = 0 \end{aligned}$$

$$\begin{aligned} &(1-\pi) \frac{\partial x_2^{h}}{\partial a^{h}} dp + (p-c)(1-\pi) \frac{\partial^2 x_2^{h}}{\partial a^{h} \partial B^{h}} dB^{h} + (p-c)(1-\pi) \frac{\partial^2 x_2^{h}}{\partial a^{h} \partial a^{h}} \\ &+ (p-c)(1-\pi) \frac{\partial^2 x_2^{h}}{\partial a^{h} \partial s} ds + (p-c)(1-\pi) \frac{\partial^2 x_2^{h}}{\partial a^{h} \partial q} (dp + dt) - dp_a = 0 \end{aligned}$$

$$
\begin{aligned}
&\pi \frac{\partial x_2^\ell}{\partial a^\ell} dp + (p-c)\,\pi \frac{\partial^2 x_2^\ell}{\partial a^\ell \partial B^\ell} dB^h + (p-c)\,\pi \frac{\partial^2 x_2^\ell}{\partial a^\ell \partial a^\ell} \\
&+ (p-c)\,\pi \frac{\partial^2 x_2^\ell}{\partial a^\ell \partial s} ds + (p-c)\,\pi \frac{\partial^2 x_2^\ell}{\partial a^\ell \partial q} (dp + dt) - dp_a = 0
\end{aligned}
$$

In matrix form we have:

$$
\begin{bmatrix}
2\left[\pi \frac{\partial x_2^\ell}{\partial q} + (1-\pi) \frac{\partial x_2^h}{\partial q}\right] + (p-c)\left[\pi \frac{\partial^2 x_2^\ell}{\partial q \partial q} + (1-\pi) \frac{\partial^2 x_2^h}{\partial q \partial q}\right] & (1-\pi)\left[\frac{\partial x_2^h}{\partial a^h} + (p-c) \frac{\partial^2 x_2^h}{\partial q \partial a^h}\right] & \pi\left[\frac{\partial x_2^\ell}{\partial a^\ell} + (p-c) \frac{\partial^2 x_2^\ell}{\partial q \partial a^\ell}\right] \\
(1-\pi)\left[\frac{\partial x_2^h}{\partial a^h} + (p-c) \frac{\partial^2 x_2^h}{\partial a^h \partial q}\right] & (p-c)(1-\pi) \frac{\partial^2 x_2^h}{\partial a^h \partial a^h} & 0 \\
\pi\left[\frac{\partial x_2^\ell}{\partial a^\ell} + (p-c) \frac{\partial^2 x_2^\ell}{\partial q \partial a^\ell}\right] & 0 & (p-c)\,\pi \frac{\partial^2 x_2^\ell}{\partial a^\ell \partial a^\ell}
\end{bmatrix}
$$

$$
\begin{bmatrix} dp \\ da^h \\ da^\ell \end{bmatrix}
=
\begin{bmatrix}
\left\{
\begin{array}{l}
-\pi\left[\frac{\partial x_2^\ell}{\partial B^\ell} + (p-c) \frac{\partial^2 x_2^\ell}{\partial q \partial B^\ell}\right] dB^\ell - (1-\pi)\left[\frac{\partial x_2^h}{\partial B^h} + (p-c) \frac{\partial^2 x_2^h}{\partial q \partial B^h}\right] dB^h \\
-\left[\left(\frac{\partial x_2^\ell}{\partial s} + (p-c) \frac{\partial^2 x_2^\ell}{\partial q \partial s}\right)\pi + \left((p-c) \frac{\partial^2 x_2^h}{\partial q \partial s} + \frac{\partial x_2^h}{\partial s}\right)(1-\pi)\right] ds \\
-\left[\left(\frac{\partial x_2^\ell}{\partial q} + (p-c) \frac{\partial^2 x_2^\ell}{\partial q \partial q}\right)\pi + \left((p-c) \frac{\partial^2 x_2^h}{\partial q \partial q} + \frac{\partial x_2^h}{\partial q}\right)(1-\pi)\right] dt
\end{array}
\right\} \\
-(p-c)(1-\pi)\left[\frac{\partial^2 x_2^h}{\partial a \partial B^h} dB^h + \frac{\partial^2 x_2^h}{\partial a^h \partial q} dt + \frac{\partial^2 x_2^h}{\partial a^h \partial s} ds\right] + dp_a \\
-(p-c)\,\pi\left[\frac{\partial^2 x_2^\ell}{\partial a \partial B^\ell} dB^\ell + \frac{\partial^2 x_2^\ell}{\partial a \ell \partial q} dt + \frac{\partial^2 x_2^\ell}{\partial a \ell \partial s} ds\right] + dp_a
\end{bmatrix}
$$

Define H^M as

$$
H^M \equiv \textit{Hessian matrix} < 0
$$

We therefore get the following comparative statics results:

$$
\begin{aligned}
&\frac{dp}{dt} \mid_{\substack{ds=0 \\ dp_a=0}} \lesseqgtr 0 \\
&\frac{da^h}{dt} \mid_{\substack{ds=0 \\ dp_a=0}} \lesseqgtr 0 \\
&\frac{da^\ell}{dt} \mid_{\substack{ds=0 \\ dp_a=0}} \lesseqgtr 0 \\
&\frac{dp}{dp_a} < 0\,,\ \frac{da^h}{dp_a} < 0\,,\ \frac{da^\ell}{dp_a} < 0 \\
&\frac{dp}{ds} > 0\,,\ \frac{da^h}{ds} > 0\,,\ \frac{da^\ell}{ds} > 0
\end{aligned}
$$

It is possible to show that only the impact on variables p, a^h and a^ℓ holding s and p_a fixed is ambiguous, while all other effects are well-defined.

Then totally differentiating the set of first order condition of platform gives:

$$2\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)dp_a+\left(\frac{\partial a^h}{\partial s}+\frac{\partial a^\ell}{\partial s}\right)ds+\left(\frac{da^h}{dt}\mid_{\substack{ds=0\\dp_a=0}}+\frac{da^\ell}{dt}\mid_{\substack{ds=0\\dp_a=0}}\right)dt+(p_a-g_{11}'')\left(\frac{\partial^2 a^h}{\partial p_a\partial s}+\frac{\partial^2 a^\ell}{\partial p_a\partial s}\right)ds$$
$$+(p_a-g_{11}'')\left[\left(\frac{\partial^2 a^h}{\partial p_a\partial p_a}+\frac{\partial^2 a^\ell}{\partial p_a\partial p_a}\right)dp_a+\left(\frac{\partial a^h}{\partial p_a}\frac{da^h}{dt}\mid_{\substack{ds=0\\dp_a=0}}+\frac{\partial a^\ell}{\partial p_a}\frac{da^\ell}{dt}\mid_{\substack{ds=0\\dp_a=0}}\right)dt\right]-g_{12}''\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)ds=0$$

$$(p_a-g_{11}'')\left(\frac{\partial^2 a^h}{\partial s\partial p_a}+\frac{\partial^2 a^\ell}{\partial s\partial p_a}\right)dp_a+(p_a-g_{12}'')\left(\frac{\partial^2 a^h}{\partial s\partial s}+\frac{\partial^2 a^\ell}{\partial s\partial s}\right)ds+\left[\frac{\partial a^h}{\partial s}+\frac{\partial a^\ell}{\partial s}-g_{12}''\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)\right]dp_a$$
$$+(p_a-g_{11}'')\left(\frac{\partial a^h}{\partial p_a}\frac{da^h}{dt}\mid_{\substack{ds=0\\dp_a=0}}+\frac{\partial a^\ell}{\partial p_a}\frac{da^\ell}{dt}\mid_{\substack{ds=0\\dp_a=0}}\right)dt-g_{ss}''ds-g_{12}''\left(\frac{da^h}{dt}\mid_{\substack{ds=0\\dp_a=0}}+\frac{da_\ell}{dt}\mid_{\substack{ds=0\\dp_a=0}}\right)dt=0$$

In matrix form we have:

$$\begin{bmatrix} 2\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)+(p_a-g_{11}'')\left(\frac{\partial^2 a^h}{\partial p_a\partial p_a}+\frac{\partial^2 a^\ell}{\partial p_a\partial p_a}\right) & (p_a-g_{11}'')\left(\frac{\partial^2 a^h}{\partial s\partial p_a}+\frac{\partial^2 a^\ell}{\partial s\partial p_a}\right)+\frac{\partial a^h}{\partial s}+\frac{\partial a^\ell}{\partial s}-g_{12}''\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right) \\ (p_a-g_{11}'')\left(\frac{\partial^2 a^h}{\partial s\partial p_a}+\frac{\partial^2 a^\ell}{\partial s\partial p_a}\right)+\frac{\partial a^h}{\partial s}+\frac{\partial a^\ell}{\partial s}-g_{12}''\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right) & (p_a-g_{12}'')\left(\frac{\partial^2 a^h}{\partial s\partial s}+\frac{\partial^2 a^\ell}{\partial s\partial s}\right)-g_{ss}'' \end{bmatrix}$$

$$\begin{bmatrix} dp_a \\ ds \end{bmatrix}=\begin{bmatrix} -\left[\frac{da^h}{dt}\mid_{\substack{ds=0\\dp_a=0}}+\frac{da^\ell}{dt}\mid_{\substack{ds=0\\dp_a=0}}+(p_a-g_{11}'')\left(\frac{\partial a^h}{\partial p_a}\frac{da^h}{dt}\mid_{\substack{ds=0\\dp_a=0}}+\frac{\partial a^\ell}{\partial p_a}\frac{da^\ell}{dt}\mid_{\substack{ds=0\\dp_a=0}}\right)\right]dt \\ -\left[(p_a-g_{11}'')\left(\frac{\partial a^h}{\partial p_a}\frac{da^h}{dt}\mid_{\substack{ds=0\\dp_a=0}}+\frac{\partial a^\ell}{\partial p_a}\frac{da^\ell}{dt}\mid_{\substack{ds=0\\dp_a=0}}\right)-g_{12}''\left(\frac{da^h}{dt}\mid_{\substack{ds=0\\dp_a=0}}+\frac{da_\ell}{dt}\mid_{\substack{ds=0\\dp_a=0}}\right)\right]dt \end{bmatrix}$$

Define H^P as

$$H^P \equiv \textit{Hessian matrix} > 0$$

Now define the term which multiplies dt in the first line as β_1 and that in the second line as β_2, then we get the following comparative statics results:

$$\begin{aligned} H^P*\frac{dp_a}{dt} &= \beta_1\left[(p_a-g_{12}'')\left(\frac{\partial^2 a^h}{\partial s\partial s}+\frac{\partial^2 a^\ell}{\partial s\partial s}\right)-g_{ss}''\right] \\ &\quad -\beta_2\left[(p_a-g_{11}'')\left(\frac{\partial^2 a^h}{\partial s\partial p_a}+\frac{\partial^2 a^\ell}{\partial s\partial p_a}\right)+\frac{\partial a^h}{\partial s}+\frac{\partial a^\ell}{\partial s}-g_{12}''\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)\right] \\ H^P*\frac{ds}{dt} &= \beta_2\left[2\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)+(p_a-g_{11}'')\left(\frac{\partial^2 a^h}{\partial p_a\partial p_a}+\frac{\partial^2 a^\ell}{\partial p_a\partial p_a}\right)\right] \\ &\quad -\beta_1\left[(p_a-g_{11}'')\left(\frac{\partial^2 a^h}{\partial s\partial p_a}+\frac{\partial^2 a^\ell}{\partial s\partial p_a}\right)+\frac{\partial a^h}{\partial s}+\frac{\partial a^\ell}{\partial s}-g_{12}''\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)\right] \end{aligned}$$

We can state that they have the opposite sign when both βs have the same sign, that is negative for βs positive and positive when they are negative, whereas their sign is ambiguous if they differ each other.

C.4 Derivation of the optimal commodity tax rate in a model with data

Start at an optimum where $t = 0$ and the nonlinear income tax is optimally chosen (i.e. the two first order conditions above apply). Then consider a tax reform that marginally increases t while at the same time adjusting B^ℓ and B^h by, respectively, $dB^\ell = x_2^\ell dt$ and $dB^h = x_2^h dt$. Consider the effects on the Lagrangian of the government's problem. We have:

$$
\begin{aligned}
&\left[\frac{\partial V^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial V^\ell}{\partial B^\ell}x_2^\ell\right]dt-\rho^\ell\frac{\partial\psi\left(x_2^\ell,a^\ell,s\right)}{\partial x_2^\ell}\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]dt\\
&+\delta\left\{\left[\frac{\partial V^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial V^h}{\partial B^h}x_2^h\right]dt-\rho^h\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial x_2^h}\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right]dt\right\}\\
&+\lambda\left[\frac{\partial V^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial V^h}{\partial B^h}x_2^h\right]dt-\lambda\rho^h\left[\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial a^h}\frac{da^h}{dt}+\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial s}\frac{ds}{dt}\right]dt\\
&-\lambda\left[\frac{\partial\widehat{V}}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial\widehat{V}}{\partial B^\ell}x_2^\ell\right]dt+\lambda\rho^h\left[\frac{\partial\psi\left(\widehat{x}_2,a^h,s\right)}{\partial a^h}\frac{da^h}{dt}+\frac{\partial\psi\left(\widehat{x}_2,a^h,s\right)}{\partial s}\frac{ds}{dt}\right]dt\\
&+\mu\left(p-c\right)\left\{\pi\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]+\left(1-\pi\right)\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right]\right\}dt\\
&+\mu\left\{\left[\pi x_2^\ell\left(B^\ell,p+t,a^\ell,s\right)+\left(1-\pi\right)x_2^h\left(B^h,p+t,a^h,s\right)\right]\frac{dp}{dt}-\left[\left(a^h+a^\ell\right)\frac{dp_a}{dt}+p_a\left(\frac{da^h}{dt}+\frac{da^\ell}{dt}\right)\right]\right\}dt
\end{aligned}
$$

Using Roy's identity and the first order conditions with respect to B^h and B^ℓ the expression above can be rewritten as:

$$\left[-x_2^\ell \frac{\partial V^\ell}{\partial B^\ell}\left(1+\frac{dp}{dt}\right)+\frac{\partial V^\ell}{\partial B^\ell}x_2^\ell\right]dt-\rho^\ell\frac{\partial\psi\left(x_2^\ell,a^\ell,s\right)}{\partial x_2^\ell}\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]dt$$
$$+\delta\left\{\left[-x_2^h\frac{\partial V^h}{\partial B^h}\left(1+\frac{dp}{dt}\right)+\frac{\partial V^h}{\partial B^h}x_2^h\right]dt-\rho^h\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial x_2^h}\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right]dt\right\}$$
$$+\lambda\left[-x_2^h\frac{\partial V^h}{\partial B^h}\left(1+\frac{dp}{dt}\right)+\frac{\partial V^h}{\partial B^h}x_2^h\right]dt-\lambda\rho^h\left[\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial a^h}\frac{da^h}{dt}+\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial s}\frac{ds}{dt}\right]dt$$
$$-\lambda\left[-\widehat{x}_2\frac{\partial\widehat{V}}{\partial B^\ell}\left(1+\frac{dp}{dt}\right)+\frac{\partial\widehat{V}}{\partial B^\ell}x_2^\ell\right]dt+\lambda\rho^h\left[\frac{\partial\psi\left(\widehat{x}_2,a^h,s\right)}{\partial a^h}\frac{da^h}{dt}+\frac{\partial\psi\left(\widehat{x}_2,a^h,s\right)}{\partial s}\frac{ds}{dt}\right]dt$$
$$+\mu\left(p-c\right)\left\{\pi\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]+\left(1-\pi\right)\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right]\right\}dt$$
$$+\mu\left\{\left[\pi x_2^\ell\left(B^\ell,p+t,a^\ell,s\right)+\left(1-\pi\right)x_2^h\left(B^h,p+t,a^h,s\right)\right]\frac{dp}{dt}-\left[\left(a^h+a^\ell\right)\frac{dp_a}{dt}+p_a\left(\frac{da^h}{dt}+\frac{da^\ell}{dt}\right)\right]\right\}dt$$

Simplifying terms we get:

$$-x_2^\ell\frac{\partial V^\ell}{\partial B^\ell}\frac{dp}{dt}dt-\left(\delta+\lambda\right)x_2^h\frac{\partial V^h}{\partial B^h}\frac{dp}{dt}dt$$
$$-\rho^\ell\frac{\partial\psi\left(x_2^\ell,a^\ell,s\right)}{\partial x_2^\ell}\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]dt$$
$$-\delta\rho^h\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial x_2^h}\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right]dt$$
$$-\lambda\rho^h\left[\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial a^h}\frac{da^h}{dt}+\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial s}\frac{ds}{dt}\right]dt$$
$$-\lambda\left[-\widehat{x}_2\frac{\partial\widehat{V}}{\partial B^\ell}\left(1+\frac{dp}{dt}\right)+\frac{\partial\widehat{V}}{\partial B^\ell}x_2^\ell\right]dt+\lambda\rho^h\left[\frac{\partial\psi\left(\widehat{x}_2,a^h,s\right)}{\partial a^h}\frac{da^h}{dt}+\frac{\partial\psi\left(\widehat{x}_2,a^h,s\right)}{\partial s}\frac{ds}{dt}\right]dt$$
$$+\mu\left(p-c\right)\left\{\pi\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]+\left(1-\pi\right)\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right]\right\}dt$$
$$+\mu\left\{\left[\pi x_2^\ell\left(B^\ell,p+t,a^\ell,s\right)+\left(1-\pi\right)x_2^h\left(B^h,p+t,a^h,s\right)\right]\frac{dp}{dt}-\left[\left(a^h+a^\ell\right)\frac{dp_a}{dt}+p_a\left(\frac{da^h}{dt}+\frac{da^\ell}{dt}\right)\right]\right\}dt$$

Using the first order conditions[3] with respect to B^h and B^ℓ the expression above can be rewritten as:

$$\begin{aligned}
&\lambda \frac{\partial \widehat{V}}{\partial B^\ell}\left[\widehat{x}_2 - x_2^\ell\right]\left(1+\frac{dp}{dt}\right)dt - \mu\left[\pi x_2^\ell + (1-\pi)\,x_2^h\right]\frac{dp}{dt}dt \\
&-\rho^\ell \frac{\partial \psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]dt \\
&-\delta\rho^h \frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right]dt \\
&-\lambda\rho^h\left[\left(\frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial a^h}-\frac{\partial \psi\left(\widehat{x}_2, a^h, s\right)}{\partial a^h}\right)\frac{da^h}{dt}+\left(\frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial s}-\frac{\partial \psi\left(\widehat{x}_2, a^h, s\right)}{\partial s}\right)\frac{ds}{dt}\right]dt \\
&+\mu(p-c)\left\{\pi\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]+(1-\pi)\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right]\right\}dt \\
&+\mu\left\{\left[\pi x_2^\ell\left(B^\ell, p+t, a^\ell, s\right)+(1-\pi)\,x_2^h\left(B^h, p+t, a^h, s\right)\right]\frac{dp}{dt}-\left[\left(a^h+a^\ell\right)\frac{dp_a}{dt}+p_a\left(\frac{da^h}{dt}+\frac{da^\ell}{dt}\right)\right]\right\}dt
\end{aligned}$$

Using the first order conditions for profit maximization of the monopolist allows us to rewrite the expression above as

$$\begin{aligned}
&\lambda \frac{\partial \widehat{V}}{\partial B^\ell}\left[\widehat{x}_2 - x_2^\ell\right]\left(1+\frac{dp}{dt}\right)dt \\
&-\rho^\ell \frac{\partial \psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]dt \\
&-\delta\rho^h \frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right]dt \\
&-\lambda\rho^h\left[\left(\frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial a^h}-\frac{\partial \psi\left(\widehat{x}_2, a^h, s\right)}{\partial a^h}\right)\frac{da^h}{dt}+\left(\frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial s}-\frac{\partial \psi\left(\widehat{x}_2, a^h, s\right)}{\partial s}\right)\frac{ds}{dt}\right]dt \\
&+\mu(p-c)\left\{\pi\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]+(1-\pi)\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right)v+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right]\right\}dt \\
&-\mu\left(a^h+a^\ell\right)\frac{dp_a}{dt}dt
\end{aligned}$$

[3]They are respectively: $(\delta+\lambda)\frac{\partial V^h}{\partial B^h} = \mu(1-\pi)$ and $\frac{\partial V^\ell}{\partial B^\ell} = \mu\pi$.

Using the first order condition of the platform we can rewrite the expression above as

$$
\begin{aligned}
&\lambda \frac{\partial \widehat{V}}{\partial B^{\ell}} \left[\widehat{x}_2 - x_2^{\ell}\right] \left(1 + \frac{dp}{dt}\right) dt \\
&-\rho^{\ell} \frac{\partial \psi\left(x_2^{\ell}, a^{\ell}, s\right)}{\partial x_2^{\ell}} \left[\frac{\partial x_2^{\ell}}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^{\ell}}{\partial a^{\ell}}\frac{da^{\ell}}{dt} + \frac{\partial x_2^{\ell}}{\partial s}\frac{ds}{dt}\right] dt \\
&-\delta\rho^{h} \frac{\partial \psi\left(x_2^{h}, a^{h}, s\right)}{\partial x_2^{h}} \left[\frac{\partial x_2^{h}}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^{h}}{\partial a^{h}}\frac{da^{h}}{dt} + \frac{\partial x_2^{h}}{\partial s}\frac{ds}{dt}\right] dt \\
&-\lambda\rho^{h} \left[\left(\frac{\partial \psi\left(x_2^{h}, a^{h}, s\right)}{\partial a^{h}} - \frac{\partial \psi\left(\widehat{x}_2, a^{h}, s\right)}{\partial a^{h}}\right)\frac{da^{h}}{dt} + \left(\frac{\partial \psi\left(x_2^{h}, a^{h}, s\right)}{\partial s} - \frac{\partial \psi\left(\widehat{x}_2, a^{h}, s\right)}{\partial s}\right)\frac{ds}{dt}\right] dt \\
&+\mu\left(p - c\right)\left\{\pi\left[\frac{\partial x_2^{\ell}}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^{\ell}}{\partial s}\frac{ds}{dt}\right] + (1-\pi)\left[\frac{\partial x_2^{h}}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^{h}}{\partial s}\frac{ds}{dt}\right]\right\} dt \\
&+\mu\left(\frac{\partial a^{h}}{\partial p_a} + \frac{\partial a^{\ell}}{\partial p_a}\right)\left(p_a - g_1'\right)\frac{dp_a}{dt} dt
\end{aligned}
$$

Or again

$$
\begin{aligned}
&\lambda \frac{\partial \widehat{V}}{\partial B^{\ell}} \left[\widehat{x}_2 - x_2^{\ell}\right] \left(1 + \frac{dp}{dt}\right) dt \\
&-\rho^{\ell} \frac{\partial \psi\left(x_2^{\ell}, a^{\ell}, s\right)}{\partial x_2^{\ell}} \left[\frac{\partial x_2^{\ell}}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^{\ell}}{\partial a^{\ell}}\frac{da^{\ell}}{dt} + \frac{\partial x_2^{\ell}}{\partial s}\frac{ds}{dt}\right] dt \\
&-\delta\rho^{h} \frac{\partial \psi\left(x_2^{h}, a^{h}, s\right)}{\partial x_2^{h}} \left[\frac{\partial x_2^{h}}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^{h}}{\partial a^{h}}\frac{da^{h}}{dt} + \frac{\partial x_2^{h}}{\partial s}\frac{ds}{dt}\right] dt \\
&-\lambda\rho^{h} \left[\left(\frac{\partial \psi\left(x_2^{h}, a^{h}, s\right)}{\partial a^{h}} - \frac{\partial \psi\left(\widehat{x}_2, a^{h}, s\right)}{\partial a^{h}}\right)\frac{da^{h}}{dt} + \left(\frac{\partial \psi\left(x_2^{h}, a^{h}, s\right)}{\partial s} - \frac{\partial \psi\left(\widehat{x}_2, a^{h}, s\right)}{\partial s}\right)\frac{ds}{dt}\right] dt \\
&+\mu\left(p - c\right)\left\{\pi\left[\frac{\partial x_2^{\ell}}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^{\ell}}{\partial s}\frac{ds}{dt}\right] + (1-\pi)\left[\frac{\partial x_2^{h}}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^{h}}{\partial s}\frac{ds}{dt}\right]\right\} dt \\
&+\mu g_2' \frac{\left(\frac{\partial a^{h}}{\partial p_a} + \frac{\partial a^{\ell}}{\partial p_a}\right)}{\left(\frac{\partial a^{h}}{\partial s} + \frac{\partial a^{\ell}}{\partial s}\right)}\frac{dp_a}{dt} dt
\end{aligned}
$$

When the reform is performed starting at a value for t which is different from zero, an additional effect needs to be considered. This effect, which stems from the fact that the reform exerts an effect on the commodity tax revenue collected by the government, is given by:

$$\mu t \left\{ \pi \left[\frac{\partial x_2^\ell}{\partial q} \left(1 + \frac{dp}{dt} \right) + \frac{\partial x_2^\ell}{\partial a^\ell} \frac{da^\ell}{dt} + \frac{\partial x_2^\ell}{\partial s} \frac{ds}{dt} \right] + (1-\pi) \left[\frac{\partial x_2^h}{\partial q} \left(1 + \frac{dp}{dt} \right) + \frac{\partial x_2^h}{\partial a^h} \frac{da^h}{dt} + \frac{\partial x_2^h}{\partial s} \frac{ds}{dt} \right] \right\}$$

Therefore, when t is optimally chosen, the following equation will hold:

$$\begin{aligned}
&\mu t \left\{ \pi \left[\frac{\partial x_2^\ell}{\partial q} \left(1 + \frac{dp}{dt} \right) + \frac{\partial x_2^\ell}{\partial a^\ell} \frac{da^\ell}{dt} + \frac{\partial x_2^\ell}{\partial s} \frac{ds}{dt} \right] + (1-\pi) \left[\frac{\partial x_2^h}{\partial q} \left(1 + \frac{dp}{dt} \right) + \frac{\partial x_2^h}{\partial a^h} \frac{da^h}{dt} + \frac{\partial x_2^h}{\partial s} \frac{ds}{dt} \right] \right\} \\
&= \\
&\lambda \frac{\partial \widehat{V}}{\partial B^\ell} \left(x_2^\ell - \widehat{x}_2 \right) \left(1 + \frac{dp}{dt} \right) \\
&+\rho^\ell \frac{\partial \psi \left(x_2^\ell, a^\ell, s \right)}{\partial x_2^\ell} \left[\frac{\partial x_2^\ell}{\partial q} \left(1 + \frac{dp}{dt} \right) + \frac{\partial x_2^\ell}{\partial a^\ell} \frac{da^\ell}{dt} + \frac{\partial x_2^\ell}{\partial s} \frac{ds}{dt} \right] \\
&+\delta \rho^h \frac{\partial \psi \left(x_2^h, a^h, s \right)}{\partial x_2^h} \left[\frac{\partial x_2^h}{\partial q} \left(1 + \frac{dp}{dt} \right) + \frac{\partial x_2^h}{\partial a^h} \frac{da^h}{dt} + \frac{\partial x_2^h}{\partial s} \frac{ds}{dt} \right] \\
&+\lambda \rho^h \left[\left(\frac{\partial \psi \left(x_2^h, a^h, s \right)}{\partial a^h} - \frac{\partial \psi \left(\widehat{x}_2, a^h, s \right)}{\partial a^h} \right) \frac{da^h}{dt} + \left(\frac{\partial \psi \left(x_2^h, a^h, s \right)}{\partial s} - \frac{\partial \psi \left(\widehat{x}_2, a^h, s \right)}{\partial s} \right) \frac{ds}{dt} \right] \\
&-\mu \left(p - c \right) \left\{ \pi \left[\frac{\partial x_2^\ell}{\partial q} \left(1 + \frac{dp}{dt} \right) + \frac{\partial x_2^\ell}{\partial s} \frac{ds}{dt} \right] + (1-\pi) \left[\frac{\partial x_2^h}{\partial q} \left(1 + \frac{dp}{dt} \right) + \frac{\partial x_2^h}{\partial s} \frac{ds}{dt} \right] \right\} \\
&-\mu \left(\frac{\partial a^h}{\partial p_a} + \frac{\partial a^\ell}{\partial p_a} \right) \left(p_a - g_1' \right) \frac{dp_a}{dt}
\end{aligned}$$

Defining Ξ as

$$\begin{aligned}
\Xi &\equiv \pi \left[\frac{\partial x_2^\ell}{\partial q} \left(1 + \frac{dp}{dt} \right) + \frac{\partial x_2^\ell}{\partial a^\ell} \frac{da^\ell}{dt} + \frac{\partial x_2^\ell}{\partial s} \frac{ds}{dt} \right] + (1-\pi) \left[\frac{\partial x_2^h}{\partial q} \left(1 + \frac{dp}{dt} \right) + \frac{\partial x_2^h}{\partial a^h} \frac{da^h}{dt} + \frac{\partial x_2^h}{\partial s} \frac{ds}{dt} \right] \\
&= \left[\pi \frac{\partial x_2^\ell}{\partial q} + (1-\pi) \frac{\partial x_2^h}{\partial q} \right] \left(1 + \frac{dp}{dt} \right) + \left[\pi \frac{\partial x_2^\ell}{\partial s} + (1-\pi) \frac{\partial x_2^h}{\partial s} \right] \frac{ds}{dt} + \pi \frac{\partial x_2^\ell}{\partial a^\ell} \frac{da^\ell}{dt} + (1-\pi) \frac{\partial x_2^h}{\partial a^h} \frac{da^h}{dt},
\end{aligned}$$

we have (taking into account that, due to the assumed quasi-linear specification of individual preferences, $\partial \widehat{V} / \partial B^\ell = 1$, $\frac{\partial \psi \left(x_2^h, a^h, s \right)}{\partial a^h} = \frac{\partial \psi \left(\widehat{x}_2, a^h, s \right)}{\partial a^h}$ and $\frac{\partial \psi \left(x_2^h, a^h, s \right)}{\partial s} =$

$\frac{\partial\psi\left(\widehat{x}_2,a^h,s\right)}{\partial s}$, since the amount of consumed commodity is the same):

$$\begin{aligned}
t &= \frac{\lambda}{\mu\Xi}\left(x_2^\ell-\widehat{x}_2\right)\left(1+\frac{dp}{dt}\right)\\
&+\frac{\rho^\ell}{\mu\Xi}\frac{\partial\psi\left(x_2^\ell,a^\ell,s\right)}{\partial x_2^\ell}\left[\frac{\partial x_2^\ell}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]\\
&+\frac{\delta\rho^h}{\mu\Xi}\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial x_2^h}\left[\frac{\partial x_2^h}{\partial q}\left(1+\frac{dp}{dt}\right)+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right]\\
&-\frac{p-c}{\Xi}\left\{\left[\pi\frac{\partial x_2^\ell}{\partial q}+(1-\pi)\frac{\partial x_2^h}{\partial q}\right]\left(1+\frac{dp}{dt}\right)+\left[\pi\frac{\partial x_2^\ell}{\partial s}+(1-\pi)\frac{\partial x_2^h}{\partial s}\right]\frac{ds}{dt}\right\}-\frac{p_a-g_1'}{\Xi}\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)\frac{dp_a}{dt}.
\end{aligned}$$

C.5 Derivation of the optimal advertising tax rate in a model with data exploitation

Start at an optimum where $t_a = 0$ and the nonlinear income tax is optimally chosen (i.e. the two first order conditions above apply). Then consider a tax reform that marginally increases t_a. Now we do not need to adjust B^ℓ and B^h. Then, consider the effects on the Lagrangian of the government's problem. We have:

$$\begin{aligned}
&\frac{\partial V^\ell}{\partial p}\frac{dp}{dt_a}dt_a-\rho^\ell\frac{\partial\psi\left(x_2^\ell,a^\ell,s\right)}{\partial x_2^\ell}\left[\frac{\partial x_2^\ell}{\partial p}\frac{dp}{dt_a}+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt_a}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt_a}\right]dt_a\\
&+\delta\left\{\frac{\partial V^h}{\partial p}\frac{dp}{dt_a}dt_a-\rho^h\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial x_2^h}\left[\frac{\partial x_2^h}{\partial p}\frac{dp}{dt_a}+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt_a}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt_a}\right]dt_a\right\}\\
&+\lambda\frac{\partial V^h}{\partial p}\frac{dp}{dt_a}dt_a-\lambda\rho^h\left[\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial a^h}\frac{da^h}{dt_a}+\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial s}\frac{ds}{dt_a}\right]dt_a\\
&-\lambda\frac{\partial\widehat{V}}{\partial p}\frac{dp}{dt_a}dt_a+\lambda\rho^h\left[\frac{\partial\psi\left(\widehat{x}_2,a^h,s\right)}{\partial a^h}\frac{da^h}{dt_a}+\frac{\partial\psi\left(\widehat{x}_2,a^h,s\right)}{\partial s}\frac{ds}{dt_a}\right]dt_a\\
&+\mu\left(p-c\right)\left\{\pi\left[\frac{\partial x_2^\ell}{\partial p}\frac{dp}{dt_a}+\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt_a}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt_a}\right]+(1-\pi)\left[\frac{\partial x_2^h}{\partial p}\frac{dp}{dt_a}+\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt_a}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt_a}\right]\right\}dt_a\\
&+\mu\left\{\left[\pi x_2^\ell+(1-\pi)x_2^h\right]\frac{dp}{dt_a}-\left[\left(a^h+a^\ell\right)\frac{dp_a}{dt_a}+p_a\left(\frac{da^h}{dt_a}+\frac{da^\ell}{dt_a}\right)\right]\right\}dt_a
\end{aligned}$$

Using Roy's identity and the first order conditions with respect to B^h and B^ℓ the expression above can be rewritten as:

$$
\begin{aligned}
&-x_2^\ell \frac{\partial V^\ell}{\partial B^\ell}\frac{dp}{dt_a}dt_a - \rho^\ell \frac{\partial \psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}\left[\frac{\partial x_2^\ell}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt_a} + \frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt_a}\right]dt_a \\
&+\delta\left\{-x_2^h \frac{\partial V^h}{\partial B^h}\frac{dp}{dt_a}dt_a - \rho^h \frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}\left[\frac{\partial x_2^h}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt_a} + \frac{\partial x_2^h}{\partial s}\frac{ds}{dt_a}\right]dt_a\right\} \\
&-\lambda x_2^h \frac{\partial V^h}{\partial B^h}\frac{dp}{dt_a}dt_a - \lambda\rho^h\left[\frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial a^h}\frac{da^h}{dt_a} + \frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial s}\frac{ds}{dt_a}\right]dt_a \\
&+\lambda \widehat{x}_2 \frac{\partial \widehat{V}}{\partial B^\ell}\frac{dp}{dt_a}dt_a + \lambda\rho^h\left[\frac{\partial \psi\left(\widehat{x}_2, a^h, s\right)}{\partial a^h}\frac{da^h}{dt_a} + \frac{\partial \psi\left(\widehat{x}_2, a^h, s\right)}{\partial s}\frac{ds}{dt_a}\right]dt_a \\
&+\mu\left(p-c\right)\left\{\pi\left[\frac{\partial x_2^\ell}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt_a} + \frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt_a}\right] + \left(1-\pi\right)\left[\frac{\partial x_2^h}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt_a} + \frac{\partial x_2^h}{\partial s}\frac{ds}{dt_a}\right]\right\}dt_a \\
&+\mu\left\{\left[\pi x_2^\ell + \left(1-\pi\right)x_2^h\right]\frac{dp}{dt_a} - \left[\left(a^h + a^\ell\right)\frac{dp_a}{dt_a} + p_a\left(\frac{da^h}{dt_a} + \frac{da^\ell}{dt_a}\right)\right]\right\}dt_a
\end{aligned}
$$

Using the first order conditions for monopolist, the expression above can be rewritten as:

$$
\begin{aligned}
&\lambda \frac{\partial \widehat{V}}{\partial B^\ell}\left[\widehat{x}_2 - x_2^\ell\right]\frac{dp}{dt_a}dt_a - \rho^\ell \frac{\partial \psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}\left[\frac{\partial x_2^\ell}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt_a} + \frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt_a}\right]dt_a \\
&-\delta\rho^h \frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}\left[\frac{\partial x_2^h}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt_a} + \frac{\partial x_2^h}{\partial s}\frac{ds}{dt_a}\right]dt_a \\
&-\lambda\rho^h\left[\left(\frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial a^h} - \frac{\partial \psi\left(\widehat{x}_2, a^h, s\right)}{\partial a^h}\right)\frac{da^h}{dt_a} + \left(\frac{\partial \psi\left(x_2^h, a^h, s\right)}{\partial s} - \frac{\partial \psi\left(\widehat{x}_2, a^h, s\right)}{\partial s}\right)\frac{ds}{dt_a}\right]dt_a \\
&+\mu\left(p-c\right)\left\{\pi\left[\frac{\partial x_2^\ell}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt_a}\right] + \left(1-\pi\right)\left[\frac{\partial x_2^h}{\partial p}\frac{dp}{dt_a} + \frac{\partial x_2^h}{\partial s}\frac{ds}{dt_a}\right]\right\}dt_a \\
&-\mu\left(a^h + a^\ell\right)\frac{dp_a}{dt_a}dt_a + \mu t_a\left(\frac{\partial a^h}{\partial t_a} + \frac{\partial a^\ell}{\partial t_a}\right)
\end{aligned}
$$

Using the first order condition of the platform we can rewrite the expression above as

$$\lambda \frac{\partial \widehat{V}}{\partial B^{\ell}} \left[\widehat{x}_2 - x_2^{\ell}\right] \frac{dp}{dt_a} dt_a - \rho^{\ell} \frac{\partial \psi \left(x_2^{\ell}, a^{\ell}, s\right)}{\partial x_2^{\ell}} \left[\frac{\partial x_2^{\ell}}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^{\ell}}{\partial a^{\ell}} \frac{da^{\ell}}{dt_a} + \frac{\partial x_2^{\ell}}{\partial s} \frac{ds}{dt_a}\right] dt_a$$
$$-\delta \rho^{h} \frac{\partial \psi \left(x_2^{h}, a^{h}, s\right)}{\partial x_2^{h}} \left[\frac{\partial x_2^{h}}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^{h}}{\partial a^{h}} \frac{da^{h}}{dt_a} + \frac{\partial x_2^{h}}{\partial s} \frac{ds}{dt_a}\right] dt_a$$
$$-\lambda \rho^{h} \left[\left(\frac{\partial \psi \left(x_2^{h}, a^{h}, s\right)}{\partial a^{h}} - \frac{\partial \psi \left(\widehat{x}_2, a^{h}, s\right)}{\partial a^{h}}\right) \frac{da^{h}}{dt_a} + \left(\frac{\partial \psi \left(x_2^{h}, a^{h}, s\right)}{\partial s} - \frac{\partial \psi \left(\widehat{x}_2, a^{h}, s\right)}{\partial s}\right) \frac{ds}{dt_a}\right] dt_a$$
$$+\mu \left(p - c\right) \left\{\pi \left[\frac{\partial x_2^{\ell}}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^{\ell}}{\partial s} \frac{ds}{dt_a}\right] + (1 - \pi) \left[\frac{\partial x_2^{h}}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^{h}}{\partial s} \frac{ds}{dt_a}\right]\right\} dt_a$$
$$+\mu \left(\frac{\partial a^{h}}{\partial p_a} + \frac{\partial a^{\ell}}{\partial p_a}\right) \left(p_a - g_1'\right) \frac{dp_a}{dt_a} dt_a + \mu t_a \left(\frac{\partial a^{h}}{\partial t_a} + \frac{\partial a^{\ell}}{\partial t_a}\right)$$

Or again

$$\lambda \frac{\partial \widehat{V}}{\partial B^{\ell}} \left[\widehat{x}_2 - x_2^{\ell}\right] \frac{dp}{dt_a} dt_a - \rho^{\ell} \frac{\partial \psi \left(x_2^{\ell}, a^{\ell}, s\right)}{\partial x_2^{\ell}} \left[\frac{\partial x_2^{\ell}}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^{\ell}}{\partial a^{\ell}} \frac{da^{\ell}}{dt_a} + \frac{\partial x_2^{\ell}}{\partial s} \frac{ds}{dt_a}\right] dt_a$$
$$-\delta \rho^{h} \frac{\partial \psi \left(x_2^{h}, a^{h}, s\right)}{\partial x_2^{h}} \left[\frac{\partial x_2^{h}}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^{h}}{\partial a^{h}} \frac{da^{h}}{dt_a} + \frac{\partial x_2^{h}}{\partial s} \frac{ds}{dt_a}\right] dt_a$$
$$-\lambda \rho^{h} \left[\left(\frac{\partial \psi \left(x_2^{h}, a^{h}, s\right)}{\partial a^{h}} - \frac{\partial \psi \left(\widehat{x}_2, a^{h}, s\right)}{\partial a^{h}}\right) \frac{da^{h}}{dt_a} + \left(\frac{\partial \psi \left(x_2^{h}, a^{h}, s\right)}{\partial s} - \frac{\partial \psi \left(\widehat{x}_2, a^{h}, s\right)}{\partial s}\right) \frac{ds}{dt_a}\right] dt_a$$
$$+\mu \left(p - c\right) \left\{\pi \left[\frac{\partial x_2^{\ell}}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^{\ell}}{\partial s} \frac{ds}{dt_a}\right] + (1 - \pi) \left[\frac{\partial x_2^{h}}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^{h}}{\partial s} \frac{ds}{dt_a}\right]\right\} dt_a$$
$$+\mu g_2' \frac{\left(\frac{\partial a^{h}}{\partial p_a} + \frac{\partial a^{\ell}}{\partial p_a}\right)}{\left(\frac{\partial a^{h}}{\partial s} + \frac{\partial a^{\ell}}{\partial s}\right)} \frac{dp_a}{dt_a} dt_a + \mu t_a \left(\frac{\partial a^{h}}{\partial t_a} + \frac{\partial a^{\ell}}{\partial t_a}\right)$$

Therefore, when t_a is optimally chosen, the following equation will hold:

$$
\begin{aligned}
&\mu t_a \left(\frac{\partial a^h}{\partial t_a} + \frac{\partial a^\ell}{\partial t_a} \right) = \lambda \frac{\partial \widehat{V}}{\partial B^\ell} \left[x_2^\ell - \widehat{x}_2 \right] \frac{dp}{dt_a} + \rho^\ell \frac{\partial \psi \left(x_2^\ell, a^\ell, s \right)}{\partial x_2^\ell} \left[\frac{\partial x_2^\ell}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^\ell}{\partial a^\ell} \frac{da^\ell}{dt_a} + \frac{\partial x_2^\ell}{\partial s} \frac{ds}{dt_a} \right] \\
&+ \delta \rho^h \frac{\partial \psi \left(x_2^h, a^h, s \right)}{\partial x_2^h} \left[\frac{\partial x_2^h}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^h}{\partial a^h} \frac{da^h}{dt_a} + \frac{\partial x_2^h}{\partial s} \frac{ds}{dt_a} \right] \\
&+ \lambda \rho^h \left[\left(\frac{\partial \psi \left(x_2^h, a^h, s \right)}{\partial a^h} - \frac{\partial \psi \left(\widehat{x}_2, a^h, s \right)}{\partial a^h} \right) \frac{da^h}{dt_a} + \left(\frac{\partial \psi \left(x_2^h, a^h, s \right)}{\partial s} - \frac{\partial \psi \left(\widehat{x}_2, a^h, s \right)}{\partial s} \right) \frac{ds}{dt_a} \right] \\
&- \mu \left(p - c \right) \left\{ \pi \left[\frac{\partial x_2^\ell}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^\ell}{\partial s} \frac{ds}{dt_a} \right] + (1 - \pi) \left[\frac{\partial x_2^h}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^h}{\partial s} \frac{ds}{dt_a} \right] \right\} \\
&- \mu \left(\frac{\partial a^h}{\partial p_a} + \frac{\partial a^\ell}{\partial p_a} \right) \left(p_a - g_1' \right) \frac{dp_a}{dt_a}
\end{aligned}
$$

Defining Ψ as

$$
\Psi \equiv \left(\frac{\partial a^h}{\partial t_a} + \frac{\partial a^\ell}{\partial t_a} \right),
$$

we have (taking into account that, due to the assumed quasi-linear specification of individual preferences, $\partial \widehat{V} / \partial B^\ell = 1$, $\frac{\partial \psi \left(x_2^h, a^h, s \right)}{\partial a^h} = \frac{\partial \psi \left(\widehat{x}_2, a^h, s \right)}{\partial a^h}$ and $\frac{\partial \psi \left(x_2^h, a^h, s \right)}{\partial s} = \frac{\partial \psi \left(\widehat{x}_2, a^h, s \right)}{\partial s}$, since the amount of consumed commodity is the same):

$$
\begin{aligned}
t_a = \; & \frac{\lambda}{\mu \Psi} \left(x_2^\ell - \widehat{x}_2 \right) \frac{dp}{dt_a} \\
& + \frac{\rho^\ell}{\mu \Psi} \frac{\partial \psi \left(x_2^\ell, a^\ell, s \right)}{\partial x_2^\ell} \left[\frac{\partial x_2^\ell}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^\ell}{\partial a^\ell} \frac{da^\ell}{dt} + \frac{\partial x_2^\ell}{\partial s} \frac{ds}{dt} \right] \\
& + \frac{\delta \rho^h}{\mu \Psi} \frac{\partial \psi \left(x_2^h, a^h, s \right)}{\partial x_2^h} \left[\frac{\partial x_2^h}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^h}{\partial a^h} \frac{da^h}{dt} + \frac{\partial x_2^h}{\partial s} \frac{ds}{dt} \right] \\
& - \frac{p - c}{\Psi} \left[\pi \left(\frac{\partial x_2^\ell}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^\ell}{\partial s} \frac{ds}{dt_a} \right) + (1 - \pi) \left(\frac{\partial x_2^h}{\partial p} \frac{dp}{dt_a} + \frac{\partial x_2^h}{\partial s} \frac{ds}{dt_a} \right) \right] - \frac{p_a - g_1'}{\Psi} \left(\frac{\partial a^h}{\partial p_a} + \frac{\partial a^\ell}{\partial p_a} \right) \frac{dp_a}{dt_a}.
\end{aligned}
$$

C.6 Comparative statics for platform and monopolist with a tax on ads

Here we show the expression needed to understand the effect of a tax on commodity through comparative statics. Totally differentiating the set of first order condition of monopolist gives:

$$\pi\left(\frac{\partial x_2^\ell}{\partial B^\ell}dB^\ell+\frac{\partial x_2^\ell}{\partial p}dp+\frac{\partial x_2^\ell}{\partial a^\ell}da^\ell+\frac{\partial x_2^\ell}{\partial s}ds\right)+(1-\pi)\left(\frac{\partial x_2^h}{\partial B^h}dB^h+\frac{\partial x_2^h}{\partial p}dp+\frac{\partial x_2^h}{\partial a^h}da^h+\frac{\partial x_2^h}{\partial s}ds\right)$$
$$+\left[\pi\frac{\partial x_2^\ell}{\partial p}+(1-\pi)\frac{\partial x_2^h}{\partial p}\right]dp+(p-c)\,\pi\left(\frac{\partial^2 x_2^\ell}{\partial p\partial B^\ell}dB^\ell+\frac{\partial^2 x_2^\ell}{\partial p\partial p}dp+\frac{\partial^2 x_2^\ell}{\partial p\partial a^\ell}da^\ell+\frac{\partial^2 x_2^\ell}{\partial p\partial s}ds\right)$$
$$+(p-c)(1-\pi)\left[\frac{\partial^2 x_2^h}{\partial p\partial B^h}dB^h+\frac{\partial^2 x_2^h}{\partial p\partial p}dp+\frac{\partial^2 x_2^h}{\partial p\partial a^h}da^h+\frac{\partial^2 x_2^h}{\partial p\partial s}ds\right]=0$$

$$(1-\pi)\left[\frac{\partial x_2^h}{\partial a^h}dp+(p-c)\left(\frac{\partial^2 x_2^h}{\partial a^h\partial B^h}dB^h+\frac{\partial^2 x_2^h}{\partial a^h\partial p}dp+\frac{\partial^2 x_2^h}{\partial a^h\partial a^h}da^h+\frac{\partial^2 x_2^h}{\partial a^h\partial s}ds\right)\right]-dp_a-dt_a=0$$

$$\pi\left[\frac{\partial x_2^\ell}{\partial a^\ell}dp+(p-c)\left(\frac{\partial^2 x_2^\ell}{\partial a^\ell\partial B^\ell}dB^h+\frac{\partial^2 x_2^\ell}{\partial a^\ell\partial p}dp+\frac{\partial^2 x_2^\ell}{\partial a^\ell\partial a^\ell}+\frac{\partial^2 x_2^\ell}{\partial a^\ell\partial s}ds\right)\right]-dp_a-dt_a=0$$

In matrix form we have:

$$\begin{bmatrix} 2\left[\pi\frac{\partial x_2^\ell}{\partial p}+(1-\pi)\frac{\partial x_2^h}{\partial p}\right]+(p-c)\left[\pi\frac{\partial^2 x_2^\ell}{\partial p\partial p}+(1-\pi)\frac{\partial^2 x_2^h}{\partial p\partial p}\right] & (1-\pi)\left[\frac{\partial x_2^h}{\partial a^h}+(p-c)\frac{\partial^2 x_2^h}{\partial p\partial a^h}\right] & \pi\left[\frac{\partial x_2^\ell}{\partial a^\ell}+(p-c)\frac{\partial^2 x_2^\ell}{\partial p\partial a^\ell}\right] \\ (1-\pi)\left[\frac{\partial x_2^h}{\partial a^h}+(p-c)\frac{\partial^2 x_2^h}{\partial a^h\partial p}\right] & (p-c)(1-\pi)\frac{\partial^2 x_2^h}{\partial a^h\partial a^h} & 0 \\ \pi\left[\frac{\partial x_2^\ell}{\partial a^\ell}+(p-c)\frac{\partial^2 x_2^\ell}{\partial p\partial a^\ell}\right] & 0 & (p-c)\,\pi\frac{\partial^2 x_2^\ell}{\partial a^\ell\partial a^\ell} \end{bmatrix}$$
$$\begin{bmatrix} dp \\ da^h \\ da^\ell \end{bmatrix}=\begin{bmatrix} -\left[\left(\frac{\partial x_2^\ell}{\partial s}+(p-c)\frac{\partial^2 x_2^\ell}{\partial p\partial s}\right)\pi+\left((p-c)\frac{\partial^2 x_2^h}{\partial q\partial s}+\frac{\partial x_2^h}{\partial s}\right)(1-\pi)\right]ds \\ -(p-c)(1-\pi)\frac{\partial^2 x_2^h}{\partial a^h\partial s}ds+dp_a+dt_a \\ -(p-c)\,\pi\frac{\partial^2 x_2^\ell}{\partial a\ell\partial s}ds+dp_a+dt_a \end{bmatrix}$$

Define H^M as

$$H^M \equiv \textit{Hessian matrix} < 0$$

We therefore get the following comparative statics results:

$$\frac{dp}{dt_a} \mid_{\substack{ds=0\\dp_a=0}} < 0$$
$$\frac{da^h}{dt_a} \mid_{\substack{ds=0\\dp_a=0}} < 0$$
$$\frac{da^\ell}{dt_a} \mid_{\substack{ds=0\\dp_a=0}} < 0$$
$$\frac{dp}{dp_a} < 0\,,\ \frac{da^h}{dp_a} < 0\,,\ \frac{da^\ell}{dp_a} < 0$$
$$\frac{dp}{ds} > 0\,,\ \frac{da^h}{ds} > 0\,,\ \frac{da^\ell}{ds} > 0$$

Then totally differentiating the set of first order condition of platform (it is equal to the previous case) gives:

$$2\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)dp_a+\left(\frac{\partial a^h}{\partial s}+\frac{\partial a^\ell}{\partial s}\right)ds+\left(\frac{da^h}{dt}\mid_{\substack{ds=0\\dp_a=0}}+\frac{da^\ell}{dt}\mid_{\substack{ds=0\\dp_a=0}}\right)dt_a+(p_a-g''_{11})\left(\frac{\partial^2 a^h}{\partial p_a\partial s}+\frac{\partial^2 a^\ell}{\partial p_a\partial s}\right)ds$$
$$+(p_a-g''_{11})\left[\left(\frac{\partial^2 a^h}{\partial p_a\partial p_a}+\frac{\partial^2 a^\ell}{\partial p_a\partial p_a}\right)dp_a+\left(\frac{\partial a^h}{\partial p_a}\frac{da^h}{dt}\mid_{\substack{ds=0\\dp_a=0}}+\frac{\partial a^\ell}{\partial p_a}\frac{da^\ell}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}\right)dt\right]-g''_{12}\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)ds=0$$

$$(p_a-g''_{11})\left(\frac{\partial^2 a^h}{\partial s\partial p_a}+\frac{\partial^2 a^\ell}{\partial s\partial p_a}\right)dp_a+(p_a-g''_{12})\left(\frac{\partial^2 a^h}{\partial s\partial s}+\frac{\partial^2 a^\ell}{\partial s\partial s}\right)ds+\left[\frac{\partial a^h}{\partial s}+\frac{\partial a^\ell}{\partial s}-g''_{12}\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)\right]dp_a$$
$$+(p_a-g''_{11})\left(\frac{\partial a^h}{\partial p_a}\frac{da^h}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}+\frac{\partial a^\ell}{\partial p_a}\frac{da^\ell}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}\right)dt_a-g''_{ss}ds-g''_{12}\left(\frac{da^h}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}+\frac{da_\ell}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}\right)dt_a=0$$

In matrix form we have:

$$\begin{bmatrix} 2\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right)+(p_a-g''_{11})\left(\frac{\partial^2 a^h}{\partial p_a\partial p_a}+\frac{\partial^2 a^\ell}{\partial p_a\partial p_a}\right) & (p_a-g''_{11})\left(\frac{\partial^2 a^h}{\partial s\partial p_a}+\frac{\partial^2 a^\ell}{\partial s\partial p_a}\right)+\frac{\partial a^h}{\partial s}+\frac{\partial a^\ell}{\partial s}-g''_{12}\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right) \\ (p_a-g''_{11})\left(\frac{\partial^2 a^h}{\partial s\partial p_a}+\frac{\partial^2 a^\ell}{\partial s\partial p_a}\right)+\frac{\partial a^h}{\partial s}+\frac{\partial a^\ell}{\partial s}-g''_{12}\left(\frac{\partial a^h}{\partial p_a}+\frac{\partial a^\ell}{\partial p_a}\right) & (p_a-g''_{12})\left(\frac{\partial^2 a^h}{\partial s\partial s}+\frac{\partial^2 a^\ell}{\partial s\partial s}\right)-g''_{ss} \end{bmatrix}$$
$$\begin{bmatrix} dp_a \\ ds \end{bmatrix}=\begin{bmatrix} -\left[\frac{da^h}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}+\frac{da^\ell}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}+(p_a-g''_{11})\left(\frac{\partial a^h}{\partial p_a}\frac{da^h}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}+\frac{\partial a^\ell}{\partial p_a}\frac{da^\ell}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}\right)\right]dt_a \\ -\left[(p_a-g''_{11})\left(\frac{\partial a^h}{\partial p_a}\frac{da^h}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}+\frac{\partial a^\ell}{\partial p_a}\frac{da^\ell}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}\right)-g''_{12}\left(\frac{da^h}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}+\frac{da_\ell}{dt_a}\mid_{\substack{ds=0\\dp_a=0}}\right)\right]dt_a \end{bmatrix}$$

Define H^P as

$$H^P \equiv Hessian\ matrix > 0$$

Now define the term which multiplies dt_a in the first line as β_1 and that in the second line as β_2, then we get the following comparative statics results:

$$\begin{aligned}
H^P * \frac{dp_a}{dt} &= \beta_1 \left[(p_a - g''_{12}) \left(\frac{\partial^2 a^h}{\partial s \partial s} + \frac{\partial^2 a^\ell}{\partial s \partial s} \right) - g''_{ss} \right] \\
&\quad -\beta_2 \left[(p_a - g''_{11}) \left(\frac{\partial^2 a^h}{\partial s \partial p_a} + \frac{\partial^2 a^\ell}{\partial s \partial p_a} \right) + \frac{\partial a^h}{\partial s} + \frac{\partial a^\ell}{\partial s} - g''_{12} \left(\frac{\partial a^h}{\partial p_a} + \frac{\partial a^\ell}{\partial p_a} \right) \right] \\
H^P * \frac{ds}{dt} &= \beta_2 \left[2 \left(\frac{\partial a^h}{\partial p_a} + \frac{\partial a^\ell}{\partial p_a} \right) + (p_a - g''_{11}) \left(\frac{\partial^2 a^h}{\partial p_a \partial p_a} + \frac{\partial^2 a^\ell}{\partial p_a \partial p_a} \right) \right] \\
&\quad -\beta_1 \left[(p_a - g''_{11}) \left(\frac{\partial^2 a^h}{\partial s \partial p_a} + \frac{\partial^2 a^\ell}{\partial s \partial p_a} \right) + \frac{\partial a^h}{\partial s} + \frac{\partial a^\ell}{\partial s} - g''_{12} \left(\frac{\partial a^h}{\partial p_a} + \frac{\partial a^\ell}{\partial p_a} \right) \right]
\end{aligned}$$

Differently from the previous case, here the sign of β_2 is negative, while the sign of β_1 is not well-defined. Therefore, we can state that when β_1 is negative too, both price for ads and data exploitation moves positively, whereas their sign remains ambiguous if β_1 is negative and moreover they can move opposite each other.

C.7 Derivation of effects on welfare subsequent the introduction of t_a when t is optimal

Defining Γ^j as:

$$\Gamma^j \equiv -\rho^j \frac{\partial \psi \left(x_2^j, a^j, s \right)}{\partial x_2^j} \left[\frac{\partial x_2^j}{\partial a^j} \left(\frac{\partial a^j}{\partial t_a} - \frac{\partial a^j}{\partial t} \frac{\frac{dp}{dt_a}}{1 + \frac{dp}{dt}} \right) + \frac{\partial x_2^j}{\partial s} \left(\frac{\partial s}{\partial t_a} - \frac{\partial s}{\partial t} \frac{\frac{dp}{dt_a}}{1 + \frac{dp}{dt}} \right) \right],$$

we have that the total effect of the reform on the Lagrangian will be given by:

$$
\begin{aligned}
&\left(\delta\Gamma^h + \Gamma^\ell\right) / \mu \\
&+t\pi \left[\frac{\partial x_2^\ell}{\partial a^\ell}\left(\frac{da^\ell}{dt_a} - \frac{da^\ell}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right) + \frac{\partial x_2^\ell}{\partial s}\left(\frac{ds}{dt_a} - \frac{ds}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)\right] \\
&+t\left(1-\pi\right)\left[\frac{\partial x_2^h}{\partial a^h}\left(\frac{da^h}{dt_a} - \frac{da^h}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right) + \frac{\partial x_2^h}{\partial s}\left(\frac{ds}{dt_a} - \frac{ds}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)\right] \\
&-\left[\pi x_2^\ell + \left(1-\pi\right)x_2^h\right]\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}} + \left(a^h + a^\ell\right) \\
&+\left(p-c\right)\pi\left[\frac{\partial x_2^\ell}{\partial a^\ell}\left(\frac{da^\ell}{dt_a} - \frac{da^\ell}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right) + \frac{\partial x_2^\ell}{\partial s}\left(\frac{ds}{dt_a} - \frac{ds}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)\right] \\
&+\left(p-c\right)\left(1-\pi\right)\left[\frac{\partial x_2^h}{\partial a^h}\left(\frac{da^h}{dt_a} - \frac{da^h}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right) + \frac{\partial x_2^h}{\partial s}\left(\frac{ds}{dt_a} - \frac{ds}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)\right] \\
&+\left[\pi x_2^\ell + \left(1-\pi\right)x_2^h\right]\left(\frac{dp}{dt_a} - \frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\frac{dp}{dt}\right) \\
&-\left(a^h + a^\ell\right)\left(1 + \frac{dp_a}{dt_a} - \frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\frac{dp_a}{dt}\right) - p_a\left[\frac{da^\ell}{dt_a} - \frac{da^\ell}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}} + \frac{da^h}{dt_a} - \frac{da^h}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right].
\end{aligned}
$$

Simplifying terms we can rewrite the above expression as:

$$
\begin{aligned}
&\left(\delta\Gamma^h+\Gamma^\ell\right)/\mu \\
&+t\pi\left[\frac{\partial x_2^\ell}{\partial a^\ell}\left(\frac{da^\ell}{dt_a}-\frac{da^\ell}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)+\frac{\partial x_2^\ell}{\partial s}\left(\frac{ds}{dt_a}-\frac{ds}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)\right] \\
&+t\left(1-\pi\right)\left[\frac{\partial x_2^h}{\partial a^h}\left(\frac{da^h}{dt_a}-\frac{da^h}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)+\frac{\partial x_2^h}{\partial s}\left(\frac{ds}{dt_a}-\frac{ds}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)\right] \\
&+\left(p-c\right)\pi\left[\frac{\partial x_2^\ell}{\partial a^\ell}\left(\frac{da^\ell}{dt_a}-\frac{da^\ell}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)+\frac{\partial x_2^\ell}{\partial s}\left(\frac{ds}{dt_a}-\frac{ds}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)\right] \\
&+\left(p-c\right)\left(1-\pi\right)\left[\frac{\partial x_2^h}{\partial a^h}\left(\frac{da^h}{dt_a}-\frac{da^h}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)+\frac{\partial x_2^h}{\partial s}\left(\frac{ds}{dt_a}-\frac{ds}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)\right] \\
&-\left(a^h+a^\ell\right)\left(\frac{dp_a}{dt_a}-\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\frac{dp_a}{dt}\right)-p_a\left[\frac{da^\ell}{dt_a}-\frac{da^\ell}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}+\frac{da^h}{dt_a}-\frac{da^h}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right].
\end{aligned}
$$

Now define

$$
A^\ell \equiv \pi\left[\frac{\partial x_2^\ell}{\partial a^\ell}\left(\frac{da^\ell}{dt_a}-\frac{da^\ell}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)+\frac{\partial x_2^\ell}{\partial s}\left(\frac{ds}{dt_a}-\frac{ds}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)\right]=-\frac{\pi\Gamma^\ell}{\rho^\ell\frac{\partial\psi\left(x_2^\ell,a^\ell,s\right)}{\partial x_2^\ell}}, \tag{C.11}
$$

$$
A^h \equiv \left(1-\pi\right)\left[\frac{\partial x_2^h}{\partial a^h}\left(\frac{da^h}{dt_a}-\frac{da^h}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)+\frac{\partial x_2^h}{\partial s}\left(\frac{ds}{dt_a}-\frac{ds}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right)\right]=-\frac{\left(1-\pi\right)\Gamma^h}{\rho^h\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial x_2^h}}. \tag{C.12}
$$

We can then rewrite our expression for the total effect of the reform on the Lagrangian of the government's problem as follows:

$$
\begin{aligned}
&\left(q-c-\frac{\rho^\ell}{\mu\pi}\frac{\partial\psi\left(x_2^\ell,a^\ell,s\right)}{\partial x_2^\ell}\right)A^\ell+\left(q-c-\frac{\delta\rho^h}{\mu\left(1-\pi\right)}\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial x_2^h}\right)A^h \\
&-\sum_{j=\ell,h}a^j\left(\frac{dp_a}{dt_a}-\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\frac{dp_a}{dt}\right)-p_a\sum_{j=\ell,h}\left(\frac{da^j}{dt_a}-\frac{da^j}{dt}\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}\right).
\end{aligned} \tag{C.13}
$$

If (C.13) is positive, then the introduction of a tax on advertising proves to be welfare-enhancing. If it is negative, it is instead a small subsidy on advertising which is welfare-enhancing. Notice also that, when $t_a = 0$, the optimal value for t is implicitly characterized by the following condition:

$$\left(q - c - \frac{\rho^\ell}{\mu\pi}\frac{\partial\psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}\right)\left[\frac{\partial x_2^\ell}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt} + \frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]\pi +$$
$$+\left(q - c - \frac{\delta\rho^h}{(1-\pi)\mu}\frac{\partial\psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}\right)\left[\frac{\partial x_2^h}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt} + \frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right](1-\pi)$$
$$+\frac{\lambda}{\mu}\left(\widehat{x}_2 - x_2^\ell\right)\left(1 + \frac{dp}{dt}\right) = \sum_{j=\ell,h}\left(a^j\frac{dp_a}{dt} + \frac{da^j}{dt}p_a\right). \tag{C.14}$$

Now rewrite (C.13) as follows:

$$\left(q - c - \frac{\rho^\ell}{\mu\pi}\frac{\partial\psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}\right)A^\ell + \left(q - c - \frac{\delta\rho^h}{\mu(1-\pi)}\frac{\partial\psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}\right)A^h$$
$$-\sum_{j=\ell,h}\left(a^j\frac{dp_a}{dt_a} + \frac{da^j}{dt_a}p_a\right) + \sum_{j=\ell,h}\left(a^j\frac{dp_a}{dt} + \frac{da^j}{dt}p_a\right)\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}. \tag{C.15}$$

Plugging into (C.15) the value for $\sum_{j=\ell,h}\left(a^j\frac{dp_a}{dt} + \frac{da^j}{dt}p_a\right)$ provided by (C.14), one can rewrite (C.15) as follows:

$$\left(q - c - \frac{\rho^\ell}{\mu\pi}\frac{\partial\psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}\right)A^\ell + \left(q - c - \frac{\delta\rho^h}{\mu(1-\pi)}\frac{\partial\psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}\right)A^h$$
$$-\sum_{j=\ell,h}\left(a^j\frac{dp_a}{dt_a} + \frac{da^j}{dt_a}p_a\right)$$
$$+\left(q - c - \frac{\rho^\ell}{\mu\pi}\frac{\partial\psi\left(x_2^\ell, a^\ell, s\right)}{\partial x_2^\ell}\right)\left[\frac{\partial x_2^\ell}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt} + \frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt}\right]\pi\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}$$
$$+\left(q - c - \frac{\delta\rho^h}{(1-\pi)\mu}\frac{\partial\psi\left(x_2^h, a^h, s\right)}{\partial x_2^h}\right)\left[\frac{\partial x_2^h}{\partial q}\left(1 + \frac{dp}{dt}\right) + \frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt} + \frac{\partial x_2^h}{\partial s}\frac{ds}{dt}\right](1-\pi)\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}$$
$$+\frac{\lambda}{\mu}\left(\widehat{x}_2 - x_2^\ell\right)\left(1 + \frac{dp}{dt}\right)\frac{\frac{dp}{dt_a}}{1+\frac{dp}{dt}}.$$

Using the definition of A^ℓ and A^h provided by (C.11) and (C.12), and simplifying and collecting terms, the expression above boils down to:

$$\left(q-c-\frac{\rho^\ell}{\mu\pi}\frac{\partial\psi\left(x_2^\ell,a^\ell,s\right)}{\partial x_2^\ell}\right)\left(\frac{\partial x_2^\ell}{\partial a^\ell}\frac{da^\ell}{dt_a}+\frac{\partial x_2^\ell}{\partial s}\frac{ds}{dt_a}+\frac{\partial x_2^\ell}{\partial q}\frac{dp}{dt_a}\right)\pi$$
$$+\left(q-c-\frac{\delta\rho^h}{\mu\left(1-\pi\right)}\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial x_2^h}\right)\left(\frac{\partial x_2^h}{\partial a^h}\frac{da^h}{dt_a}+\frac{\partial x_2^h}{\partial s}\frac{ds}{dt_a}+\frac{\partial x_2^h}{\partial q}\frac{dp}{dt_a}\right)\left(1-\pi\right)$$
$$-\sum_{j=\ell,h}\left(a^j\frac{dp_a}{dt_a}+\frac{da^j}{dt_a}p_a\right)+\frac{\lambda}{\mu}\left(\widehat{x}_2-x_2^\ell\right)\frac{dp}{dt_a}.$$

Finally, defining $\frac{dx_2^j}{dt_a}$ as $\frac{dx_2^j}{dt_a}\equiv\frac{\partial x_2^j}{\partial a^j}\frac{da^j}{dt_a}+\frac{\partial x_2^j}{\partial s}\frac{ds}{dt_a}+\frac{\partial x_2^j}{\partial q}\frac{dp}{dt_a}$ and the elasticity ϵ_{a^j,p_a} as $\epsilon_{a^j,p_a}\equiv\frac{da^j}{dp_a}\frac{p_a}{a^j}$, we can rewrite the expression above as:

$$\left(q-c-\frac{\rho^\ell}{\mu\pi}\frac{\partial\psi\left(x_2^\ell,a^\ell,s\right)}{\partial x_2^\ell}\right)\pi\frac{dx_2^\ell}{dt_a}+\left(q-c-\frac{\delta\rho^h}{\mu\left(1-\pi\right)}\frac{\partial\psi\left(x_2^h,a^h,s\right)}{\partial x_2^h}\right)\left(1-\pi\right)\frac{dx_2^h}{dt_a}$$
$$-\sum_{j=\ell,h}\left(1+\epsilon_{a^j,p_a}\right)a^j\frac{dp_a}{dt_a}+\frac{\lambda}{\mu}\left(\widehat{x}_2-x_2^\ell\right)\frac{dp}{dt_a}.$$