UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

TESIS DOCTORAL

Regulations and strategies for the digital transformation optimal expansion of networks and provision of digital goods and services

Regulación y estrategias para la transformación digital eficiencia en la expansión de redes y en la provisión de bienes y servicios digitales

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

José Marino García García

DIRECTORES

Aurelia Valiño Castro
Antonio Jesús Sánchez Fuentes

Madrid, 2018

REGULACIÓN Y ESTRATEGIAS PARA LA TRANSFORMACIÓN DIGITAL EFICIENCIA EN LA EXPANSIÓN DE REDES Y EN LA PROVISIÓN DE BIENES Y SERVICIOS DIGITALES

REGULATIONS AND STRATEGIES FOR THE DIGITAL TRANSFORMATION OPTIMAL EXPANSION OF NETWORKS AND PROVISION OF DIGITAL GOODS AND SERVICES

UNIVERSIDAD COMPLUTENSE MADRID

TESIS DOCTORAL

José Marino García García

Directores

Aurelia Valiño Castro

Antonio Jesús Sánchez Fuentes

FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

2017

ACKNOWLEDGEMENTS

The stimulus for developing this Ph.D. thesis comes from the fascination and profound curiosity about the intersection between new digital technologies, regulation and economic analysis together with the observation on how research and professional practice on these fields are more often than not disconnected. I would like to thank everyone that supported me in this endeavor and shaped the idea of carrying out research in the junction between these areas and helped and oriented the process of writing and developing this project.

First of all, I would like to express my gratitude to my supervisors Aurelia Valiño Castro and Antonio Jesús Sánchez Fuentes for all the valuable advice and guidance throughout the elaboration of this thesis always pointing to the right direction. Without their help this research could not have been possible. I would also like to thank Aurelia for encouraging and helping me to seek for study and research abroad. This gave me the opportunity to deeply widen my knowledge and broad my perspectives for this thesis and also as an individual.

I am also extremely grateful to the Fulbright Program of the Ministry of Industry Energy and Tourism that make it possible to highly enrich this thesis through a scholarship that allowed me to pursue studies of economic policy management at Columbia University and to do research at the World Bank during a fourteen month stay in the United States.

Besides the contributions of my supervisors, this thesis has also greatly benefited from the comments and fruitful discussions with many colleagues at the World Bank, the Ministry of Industry Energy and Tourism and with professors at Complutense and Columbia University. I would like to specially thank Raul Katz for his very useful comments on Chapter 3 and Timothy John Charles Kelly for fruitful

discussions on spectrum management and on the economics, the policies and regulations suitable and available to expand digital connectivity.

Finally, but no less important, the elaboration of this thesis would have not been possible without the support and love of my family to which I dedicate the results of this work.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The digital transformation taking place in the last decades is a revolution with a deep economic, social, and cultural impact. It has created enormous opportunities and also new risks. Growth, enhanced productivity, increased citizenship engagement, better democracy and social inclusion are some of the opportunities but at the same time these transformations are challenging human and consumer rights, social policies and economic efficiency. Reaching the huge potential benefits of the digital revolution requires economic analysis to guide new regulations and policies accompanying technological advancement.

This thesis proposes regulations and strategies for an optimal provision of digital networks, goods and services contributing to some of the topics of focus in recent literature on digital economy. One problem is entry barriers to companies selling internet access. We contribute to the literature by analyzing one of the causes of the existence of entry barriers, the problem of how efficiently allocate the radio-spectrum. Another problem is the optimal prices to remunerate network providers, digital service providers, consumers and vendors. We contribute by examining the impact of price discrimination of internet content providers under duopolistic competition and multi-dimensional product differentiation in retail broadband access in chapter 3.

Chapter 1 offers some proposals for the evaluation of the efficient allocation of spectrum to radio communication services. New approaches to spectrum management have resulted in a more efficient production of services. However, the role of the public sector is still essential in spectrum allocation. This chapter provides a methodology to measure the net benefit of the reallocation of a spectrum band intended to guide regulators and policy makers. We have identified the following facts.

- The calculation of benefits and of the opportunity costs of spectrum usage should include the external value associated to the provision of services.
- Demand for spectrum is driven by the quantity and quality of services consumed. In some cases, the population density is also a driver of spectrum demand.
- The cost of deploying networks to use the spectrum is a key element to determine the private value associated to the use of a frequency band. However, there are other inputs such as the number of stations, or the use of wired solutions that may be substitutes of a frequency band.
- The utility of spectrum usage is variable. Different portions of spectrum are not perfect substitutes one another, consequently, frequencies have variable substitution margins.
- Reallocation of spectrum produces transition costs that should be included in the cost benefit analysis. Sometimes the reallocation may also reduce costs if the provision of the incumbent service is streamlined.

- Economies of scale in the manufacture of the equipment required to use a frequency band can be obtained if the band is planned to be allocated in a harmonized way across a large geographical area with high expected number of customers.

Chapter 2 describes the optimal path and speed of spectrum management reform under associated uncertain costs and benefits. This chapter offers a model to ascertain the optimal speed of reform and shed some light on the policy options to reduce uncertainty in the associated costs and benefits of reform. We have determined when to choose a gradual or big bang reform depending on technology and whether to wait or not for a new technological advancement. Both results depend on the expected outcome of the reform, the probability of realization of the considered scenarios and the value of reversal to the previous situation.

Chapter 3 presents a model of price discrimination of over the top providers under duopolistic competition and multi-dimensional product differentiation in retail broadband access. We discuss the impact of product differentiation and price discrimination on social welfare, and offer systematic simulations using feasible ranges for parameters value to help discern the impact of departing from network neutrality regulation on social welfare.

Abandoning network neutrality regulation reduces the quantity of internet content produced. Network diversity produces a significant reduction of prioritized content while the quantity of content non-susceptible of being prioritized remains unchanged. The reduction in the quantity of content is equal in the two network diversity scenarios considered – both and only one access provider prioritizing traffic.

Under network diversity access providers enjoy higher income coming from the tax charged to content providers for prioritization but content providers offering prioritized content, and advertisers suffer an income reduction. Abandoning network neutrality has two different types of effects on consumer's utility. On one hand, it reduces utility through a decrease in the amount of content. On the other hand, the disutility caused by the delay of data packages is reduced. The simulation is necessary to measure net impact. Simulations show that network neutrality regulation is welfare superior to network diversity under the model assumptions with the values and value ranges given to parameters. Departing from network neutrality regulation leads to an abrupt decrease of the consumer surplus of internet users. This effect weights more in total welfare than the increase in the access providers' surplus coming from the tax charged to content providers.

RESUMEN

La transformación digital que está teniendo lugar en las últimas décadas es una revolución con un profundo impacto económico, social y cultural que ha creado enormes oportunidades y también nuevos retos. Algunas de las oportunidades son un mayor crecimiento económico, mejoras de la productividad, mayor participación ciudadana y una mayor inclusión social, pero al mismo tiempo estas transformaciones desafían los derechos de los ciudadanos y de los consumidores, las políticas sociales y la eficiencia económica. Para alcanzar el enorme potencial de la revolución digital es necesario que el análisis económico guie la regulación y las políticas que acompañan al desarrollo tecnológico.

Esta tesis propone nuevas normas y estrategias para una óptima provisión de redes, bienes y servicios digitales, contribuyendo en algunos de los temas de interés de la literatura reciente sobre economía digital. Un problema es la existencia de barreras a la entrada en el mercado de provisión de servicios de acceso a internet. Contribuimos a la literatura existente sobre este asunto, analizando en los capítulos 1 y 2, una de las causas de la existencia dichas barreras, la atribución ineficiente del espectro radioeléctrico. Otro problema es la formación de precios socialmente óptima para remunerar a los proveedores de acceso a internet, los proveedores de servicios digitales, los usuarios y los vendedores de publicidad. Contribuimos a esta materia examinando en el capítulo 3 el impacto de la discriminación de precios sobre proveedores de contenidos cuando en el mercado minorista de acceso a internet existe competencia imperfecta, duopolio, y diferenciación multidimensional de producto.

El capítulo 1 ofrece propuestas para la evaluación de la eficiencia de una atribución de espectro a un servicio de radiocomunicaciones. Las nuevas técnicas de gestión del espectro han dado lugar a una provisión más eficiente de los servicios. Sin embargo, el papel del sector público sigue siendo necesario en la gestión del espectro. Este capítulo ofrece una metodología para medir el beneficio neto de una re-atribución de una banda de frecuencias para guiar a reguladores y diseñadores de políticas públicas. Se han obtenido los siguientes resultados destacables.

- Las externalidades asociadas a la provisión de un servicio que utiliza el espectro son relevantes y deben incluirse en el cálculo del beneficio o del coste de oportunidad de utilizar o no el espectro.

- La demanda de espectro viene determinada por la cantidad y calidad de los servicios finales que usan el espectro. En algunos casos, la densidad de población también modifica la intensidad de la demanda de espectro.

- El coste de despliegue de las redes que usan el espectro es un elemento clave para determinar el valor privado asociado al uso de una banda de frecuencias. Sin embargo, existen otros insumos, tales como el número de estaciones desplegadas, o la disponibilidad de otras tecnologías que son sustitutas de la utilización de una banda de frecuencias.

- La utilidad de las bandas de frecuencias no es uniforme. En consecuencia, las bandas de frecuencias tienen márgenes de sustitución variables.

- La re-atribución de bandas de frecuencias genera costes de transición que deben incluirse en el análisis coste beneficio. En algunas ocasiones, la re-atribución puede reducir los costes totales si se racionaliza la provisión del servicio existente.

- Pueden obtenerse economías de escala en la fabricación de equipos si la utilización de una banda de frecuencias se armoniza en un área geográfica amplia en la que haya un gran número de usuarios.

El capítulo 2 analiza el proceso de reforma de la gestión del espectro y la velocidad óptima de transformación cuando existe incertidumbre sobre los costes y beneficios de la reforma, el capítulo también arroja luz sobre las diferentes políticas para reducir la incertidumbre. Hemos determinado cuando elegir una reforma gradual y una conjunta dependiendo del estado de arte de la tecnología. También hemos proporcionado una herramienta para ayudar al regulador a tomar la decisión de adoptar inmediatamente una tecnología o esperar a su madurez. La respuesta a ambas cuestiones depende de las expectativas sobre el resultado de la reforma, la probabilidad de ocurrencia de los escenarios considerados y del coste de volver al estado anterior a la reforma.

El capítulo 3 presenta un modelo de discriminación del precio que pagan los proveedores de contenidos por acceder a internet, en un entorno de competencia en duopolio y diferenciación de producto multidimensional en el mercado minorista de acceso a internet. Discutimos el impacto de la diferenciación de producto y la discriminación de precios sobre el bienestar total, y ofrecemos simulaciones sistemáticas utilizando rangos de valores posibles para los parámetros considerados en el modelo de modo que ayuden a discernir el impacto de abandonar la regulación de neutralidad de la red sobre el bienestar total.

Abandonar la regulación de neutralidad de la red produce una reducción significativa del contenido de internet susceptible de ser priorizado, mientras que la cantidad de contenido que no es susceptible de ser priorizada permanece inalterada. La reducción en la cantidad de contenido es igual en los dos escenarios de diversidad de la red considerados en el modelo. En un caso los dos proveedores de acceso priorizan y en el otro sólo un proveedor de acceso prioriza el tráfico y el otro no.

Sin neutralidad de la red, los proveedores de acceso a Internet disfrutan de mayores beneficios gracias a la tasa por priorización que cargan a los proveedores de contenidos, pero los proveedores de contenidos dispuestos a pagar por la priorización y los anunciantes ven reducido su beneficio. El abandono de la neutralidad de la red tiene dos efectos sobre la utilidad de los consumidores. Por un lado, disminuye la utilidad como consecuencia de la reducción en la cantidad de contenidos. Por otro lado, aumenta porque disminuye el número de paquetes susceptibles de priorización. La simulación es necesaria para medir el impacto neto. La simulación muestra que la neutralidad de la red genera un mayor bienestar total para los supuestos considerados con los valores dados a los parámetros del modelo. Abandonar la neutralidad de la red produce un abrupta caída del beneficio de los consumidores. Este efecto pesa más en el bienestar total que el incremento del beneficio del productor de acceso a internet procedente de la tasa cargada a los proveedores de contenidos.

INTRODUCTION

The digital transformation taking place in the last decades is a revolution with a deep economic, social, and cultural impact. It has created enormous opportunities and also new risks. Growth, enhanced productivity, increased citizenship engagement, better democracy and social inclusion are some of the opportunities but at the same time these transformations are challenging human and consumer rights, social policies and economic efficiency. Reaching the huge potential benefits of the digital revolution requires economic analysis to guide new regulations and policies accompanying technological advancement.

The digital dimension of human and constitutional rights must be protected. Digital privacy, online freedom of expression or cyber security are examples of regulatory challenges that need to be dealt by public authorities to ensure that the digital world has at least the same level of protection of rights than the analogue. Where our data is flowing and under what conditions it is stored and manipulated are questions whose answer depend the respect for fundamental rights. The internet can be a platform to enhance freedom of expression, democracy and citizenship engagement but if institutions are not accountable and regulations are not in place it can also be used to increase information control. Thus, the internet can be a multiplier or a divider of transparency, democracy and accountability of institutions. Security has a new dimension, the protection of digital goods, private data, and the fight against cybercrime. The balance between security and privacy is now to be redefined in the digital world.

Regulators also have to cope with new forms of social exclusion. Digital technologies may be a powerful tool for inclusion but at the same time of social division. Inclusion stems from the increased availability of information and access to public and private services at lesser cost for those that are able to enter the digital society. For the others that aren´t, the internet is a new form of exclusion. Digital literacy and policies to help those that can´t afford internet access should be incorporated to social policies in order to reduce the digital divide. Besides, the digital revolution is increasing the gap in the jobs market between highly educated workers with digital skills and the less educated that usually lack of this type of training. The incorporation of academic subjects about digital technologies in education would help to reduce this gap, adapting skills to the new jobs of the digital economy.

Regulations are also necessary for ensuring economic efficiency. The digital transformation impacts the economy by creating new goods and services, reducing the cost of accessing, processing and understanding information, easing transactions and improving markets performance but if not regulated it involves the risk of reduced market competition through the creation of monopoly and monopsony power not only in the new digital goods and services markets but also extending market power to existing sectors.

The role of the public sector promoting economic efficiency in the digital age has multiple perspectives. Preventing and correcting coordination failure by favoring the adoption of standards and the interoperability of networks and services, defining proper taxation schemes in order to develop digital goods and services, providing e-government services, creating public-private partnerships to finance network deployment where the private sector either lacks of sufficient funding or incentives to deploy the optimal level of infrastructure, designing state aid programs enabling research and development of innovative technologies that otherwise would not be carried out.

Recent industrial organization literature analyzing digital economy related issues focuses on three important fields to take account of, the regulation of network deployment and access, the provision of digital services and the transactions and interactions among networks, digital services providers, consumers and vendors.

The deployment of digital networks is intrinsically involved in high entry barriers, which hinder effective competition, and suboptimal provision due to the existence of externalities. Two are the main reasons for entry barriers to exist, the economics of deployment of some network segments and the restricted access to a scarce resource, the radio-electric spectrum, which is indispensable for networks providing wireless services.

The economics of digital networks is defined by high sunk cost and scale, scope and density economies. The existence of high sunk costs hampers market entry because of the huge financial capacity required to build the network. Scale economies make the cost per user drop with higher number of users. Markets with strong scale economies tend to be less competitive due to the higher cost per user that the entrants have to face compared to the incumbents. With economies of scope the average total cost of production decreases when increasing the number of services provided. Scope economies explain why fixed and mobile broadband and phone communications are usually offered in a bundle. Companies unable to jointly provide the bundle, for example because of lack of access to spectrum, may not enter any of these services' markets. A consequence of scope economies is market concentration. This phenomenon has been a salient trend in the last decade. Economies of density make network deployment profitable in highly populated areas but cost ineffective in those sparsely populated. As spillover effects are not captured by private investors, network supply may be suboptimal without regulation. Coverage obligations, network access price regulation and mandated infrastructure sharing are several of the regulatory tools available but the design and refining of such regulations in an environment of technological evolution still needs more research to be carried out.

A second regulatory area of importance related to the regulation of networks is spectrum management. The radio-electric spectrum is the essential raw material for the provision of digital goods and services. Efficient spectrum allocation is hampered by its physical features, the nature of some services using the spectrum and inefficiencies of the allocation process. The allocation of spectrum to services has been planned by governmental agencies practically since the early days of radio communications. Governmental planning has been the response to the complex definition of property rights, high enforcement and transaction costs linked to the existence of harmful electromagnetic interferences, services that are pure public goods, externalities associated to services, and economies of scale when spectrum allocation is harmonized.

In the last decades, improved technology enabling services to share the same portion of spectrum, better understanding of interference solving and spectrum markets and especially the extraordinary growth in the demand for spectrum have put deep pressure to traditional allocation methods. As a result, the reform of spectrum has become a salient policy issue and multiple new approaches to spectrum management have emerged. Examples of spectrum management reforms are the use of auctions to grant licenses, the authorization of secondary market transactions -transfer, leasing, mutualisation of usage rights- , the use of administered incentive pricing, spectrum sharing techniques and the definition of licenses in terms of acceptable interference parameters.

Dealing with market power in the provision of digital services is another research field of interest. In this case, market power is generated in the demand side instead of supply. Utility obtained by consumers from the consumption of digital goods and services increases with the number of users using the same service. This fact results in market concentration in one or in a few providers. Additionally, transaction costs of buying goods using digital services are lower than those using traditional means. If the provision of digital goods is concentrated, market power can be extended to traditional sectors from the side of the digital demand. Both factors support the creation of monopsony power. These effects combined may generate two adverse consequences, new digital companies attempting to enter the market will face increased barriers and consumers would find higher switching costs.

Finally, the transactions and interactions among network providers, digital services providers, consumers and vendors raise difficult questions about the optimal prices to remunerate all involved parties efficiently. Most of the goods of the digital revolution are provided through two sided markets. These types of markets involve two different classes of customers, and a vendor, the platform, that enables transactions. Network providers are a platform for the transactions among internet users and internet content aggregators. At the same time, content aggregators, are a platform for the transactions among content providers and internet users. This complex structure together with reduced competition, potential monopoly power in network provision and monopsony power in the provision of services, difficult the optimal provision of digital goods and services.

This thesis proposes regulations and strategies for the economic efficiency and the optimal provision of digital networks, goods and services contributing to some of the topics of focus in recent literature. One field of interest is entry barriers to companies selling internet access. We contribute to the literature by analyzing the efficient allocation of spectrum, one of the entry barriers' causes, and the optimal path and speed of spectrum management reform. Another field on interest is the calculation of the optimal prices to remunerate network providers, digital service providers, consumers and vendors. We contribute to the literature by examining the impact of price discrimination of internet service providers under duopolistic competition and multi-dimensional product differentiation in retail broadband access.

The thesis is structured in two blocks. The first block has two chapters. It focuses on spectrum management as a key input for the supply of digital networks and consequently for the provision of digital goods and services. We first tackle the problem from a wide ranged angle, considering spectrum management as a field of steady progress and permanent change, approaching this field to the arena of the political economy analysis of reform. We expand the model of Dewatripont and Roland (1995) adding the idiosyncrasies of spectrum management and replacing political uncertainness by the uncertainty generated by unknown costs and benefits of spectrum management reform. Secondly, we analyze how to approach a specific and significant reform case, one where the regulator needs to ascertain the most efficient solution to re-allocate a highly demanded portion of spectrum with an incumbent user, new candidates to use the spectrum and associated externalities. We propose a methodology to analyze the efficiency of the allocation of spectrum to a radio-communication service in such types of spectrum bands and situations.

The second block has one extensive chapter. It discusses the optimal regulation of the price to access the internet in a context of imperfect competition, which is the common market structure. Our setup departs from the model described by Liu and Shuai (2015) with multi-dimensional product differentiation of the internet access product. One dimension is related to the set of products bundled to internet access. The other, represents how users value package delay or a delay variation that may worsen the quality of some internet services and applications. We consider the possibility of price discriminating providers of digital content when they buy internet connectivity. Our goal is to determine the price structure maximizing total welfare. Our model uses a sequential moves game theoretic approach to analyze total welfare implications of price discrimination. Finally, we offer systematic simulations, whose objective is to fill the gaps that cannot be easily explained using only mathematical analysis.

The content of each Chapter of this thesis can be briefly summarized as follows:

Chapter 1: Some proposals for the evaluation of the efficient allocation of spectrum for radio communication service. New approaches to spectrum management have resulted in more efficient production of services enabling better quality of service, a reduction of the used amount of spectrum, increased coverage and reduced prices. However, the role of the public sector is still essential in spectrum allocation. Technology advancement and improvements of the licensing methods must be complemented with processes of reallocation led by the public sector enabling higher allocative efficiency of resources. This chapter provides a methodology to measure the net benefit of the reallocation of a spectrum band intended to guide regulators and policy makers.

Chapter 2: Path and speed of spectrum management reform under associated uncertain costs and benefits. The question in spectrum management is no longer if reform is necessary to give market more participation but the path and speed of reform with current and expected technology advance. This chapter aims to contribute to answer this question offering a model to ascertain the optimal speed of reform and shed some light on the policy options to reduce uncertainty in the associated costs and benefits of reform.

In this chapter, we analyze reform types and the associated uncertainty stemming from the intrinsic physical features of spectrum and the insufficient internalization of technology advancements in regulations. We discuss how existing and upcoming technology may help overcome market failure and uncertainness.

Chapter 3: Price discrimination of over the top providers under duopolistic competition and multi-dimensional product differentiation in retail broadband access. Network neutrality regulation prevents price discrimination from access providers to content providers and product differentiation in terms of connection quality in the retail broadband access market. This chapter analyzes the economic implications of price discrimination under duopolistic competition and multi-dimensional product differentiation in retail internet access using a sequential-moves game theoretic model. Under this framework, we discuss the impact of product differentiation and price discrimination on social welfare, and offer systematic simulations using feasible ranges for parameters value to help discern the impact of departing from network neutrality regulation on social welfare.

Finally we offer some conclusions and suggest extensions and future research.

CHAPTER1: SOME PROPOSALS FOR THE EVALUATION OF THE EFFICIENCY OF AN ALLOCATION OF SPECTRUM TO RADIO COMMUNICATION SERVICES

Abstract: *New approaches to spectrum management have enabled a more efficient provision and better quality of services, a reduction of the amount of spectrum required, increased coverage and reduced prices. However, the role of the public sector is still essential in spectrum allocation and management. Technology advancement and improvements in the licensing procedures must be complemented with processes of reallocation led by the public sector enabling higher allocative efficiency of resources. This paper provides a methodology to measure the net benefit of the reallocation of a spectrum band intended to guide regulators and policy makers in the process.*

JEL Classification: D61, D62, L52, L96.

Key words: radio-electric spectrum, allocative efficiency, property rights, externalities.

1. Introduction

The radio-electric spectrum, a subset of the electromagnetic waves, is the essential raw material for the provision of a myriad of different services of increasing importance for the economy and society. Some spectrum usages are well-known and extensively consumed like radio, television, mobile voice communications and broadband. National defense, emergency assistance and scientific research are also users of spectrum. Future expected applications include telemedicine diagnosis and treatment, automated vehicles, monitoring of electricity and water networks for a sustainable consumption and efficient traffic management solutions and transportation systems. The list of services and applications is growing and spectrum use is changing the economy and society. Many of the economic and social benefits of the digital transformation taking place would not be possible without an efficient allocation of spectrum to different services.

Efficient spectrum allocation is hampered by its physical features and the nature of some services using the spectrum. Physical features have been the underpinning rationale for public management of allocation. For example, the potential existence of harmful electromagnetic interference among spectrum users obstructs proper definition and enforcement of property rights and secondary market transactions. Interferences may happen between users providing different services, intra service interference, or among providers of the same service, inter service interference. Besides, there exist services that are pure public goods, externalities associated to services and economies of scale that can only be leveraged when spectrum allocation is harmonized. However, more efficient approaches and procedures and higher market participation in spectrum management and allocation has been early and consistently proposed; Coase (1959, 1960), Levin (1970), Hazlett (1998, 2003), Kwerel and Williams (2002), Baumol and Robyn (2006).

Scarcity brought about by the huge growth in the demand for spectrum, especially in some bands demand clearly exceeds supply, has led to a consensus among regulators and stakeholders on the benefits of lighter regulation and higher flexibility in license conditions. As a result, many significant developments in spectrum management have taken place recently. Such changes are expected to reduce the cost and the speed of the deployment of networks carrying new valuable services and applications. The use of auctions to grant licenses, technology change without prior administrative authorization, and the allowance of secondary market transactions, such as transfers, leases, spectrum licenses pooling and the mutualisation of usage rights are examples of new approaches to spectrum management that are usual recommendations for inclusion on Digital National Strategies, see Broadband Commission (2014, 2016). This new possibilities in the allocation process have resulted in a more efficient production of services enabling better quality, a reduction of the amount of spectrum required to provide services, increased coverage and reduced prices. However, developments intended to achieve better allocation of services into portions of spectrum, such as spectrum sharing techniques or the definition of licenses in terms of acceptable interference parameters still has a long way to efficiently enable the market to carry out intra service transactions of usage rights. The impact of sharing is diminished because it is limited to some extent to short range applications, and acceptable interference licenses because more research is still required for and effective definition that reduces huge transaction uncertainty.

In this context, the role of the public sector as the agent coordinating the reallocation of frequencies to services with higher social value is still essential in spectrum management. New spectrum availability brought about by technology advancement, shared use of spectrum or a more efficient use of currently allocated frequencies must be complemented with processes of reallocation with well-defined and transparent transactions and reduced uncertainty for both the incumbent and the entrants to a new portion of available spectrum. The purpose of public sector intervention is to introduce certainty in transactions, ensuring the production of public goods and enabling the generation of the external value associated to services.

This paper aims to contribute to measure the social value of using a spectrum band for one or more services. Bazelon and McHenery (2013), Alden (2012), Prasad (2014) and Cave et al (2016) added to the definition of methodologies for estimating the value of spectrum taking account of externalities linked to services. While these studies analyzed how to measure social value from a static point of view, we focus on measuring the resulting value of the dynamic process of reallocating a spectrum band previously used by one or more incumbent services planned to be also used by one or more entrants.

This paper is structured as follows, section 2 analyzes current spectrum management models, section 3 describes the new models available that may increase the efficiency of spectrum use, section 4 offers a methodology to evaluate the social value of the reallocation of a portion of spectrum and section 5 concludes.

2. Spectrum management models

Spectrum management is shaped by the physical features of wave propagation and by the nature of services using the spectrum. A single management model is not possible for all frequency bands because separate bands have different features which mold the way in which frequencies are allocated and managed. Additionally, the use of spectrum is an input for the provision of services but not the only possible solution. The allocation of additional spectrum to a service may be substituted by the deployment of dense networks with a higher number of radio stations but with increased cost of network deployment.

Frequency bands are different in terms of their capacity to cover a geographical area and the amount of information that are capable to carry. In general, higher frequencies cover less extended areas but instead can transport more information. Alternatively, low frequencies help to minimize the number of stations required to provide some services because they can cover larger areas, thus reducing the cost of network deployment. Consequently, only a portion of the radio-electric spectrum has physical features enabling at the same time enough capacity to carry reasonable amount of information with limited network deployment costs. These bands are highly demanded and service providers are willing to pay more to have access to such bands.

Spectrum can be used to provide a huge variety of services. The suitability of a frequency band to provide a service depends on the nature of the service being provided. Whereas some require high transmitting power like broadcasting, others are intended to provide short ranged applications. If the service is intended to provide communication services, requirements are different for communication between persons or machines and for one-way or multidirectional. The provision of some services require the allocation of a combination of different portions of spectrum, including high frequencies in urban locations to serve densely populated areas and low frequencies in rural locations to maximize the coverage of each station if population is spread out.

Spectrum management is carried out from both, international and national public institutions. International agreements are required because without coordination between countries many spectrum usages may interfere with services of neighboring countries. For example, the frequencies used by a country to provide a television channel in a border area may not be used by a neighbor in the areas of its country near the border. Consequently, coordination of spectrum allocation and cooperation among countries ease the deployment of telecommunication networks. When neighbors use the spectrum in the same way the probability of interference is reduced and the agreements to distribute frequencies between countries in border areas are much easier to reach. Besides, coordination of allocation reduce the cost of providing services because the higher the number devices using the same portion of spectrum the higher the economies of scale in the manufacturing of the electronic communication components.

A specialized agency of the United Nations, the International Telecommunication Union (ITU), is responsible for the registration of internationally recognized allocations in the table of frequency allocations. An allocation is the registration of a particular frequency band to be used for a particular service. Additionally, there is a registration of allotments and assignments. An allotment is the right to use a particular frequency of the allocated band in a geographical area. An assignment identifies a particular radio-station defining not only the frequency being used but also a set of conditions, such the location, the transmission power and the direction of emission. The ITU is also responsible for the elaboration of the Radio Regulations, an international treaty ratified by ITU Member States, which contains inter alia rules for spectrum usage and rules of procedure to change allocations, allotments and assignments. Radio regulations are reviewed periodically in the World Radio Communication conferences in order to accommodate new services and technologies in unused bands and promoting the reallocation of existing services.

Nationally, each country can allocate the spectrum within their borders according to national preferences provided that usages differential with respect to the international allocations don't cause harmful interference to services of neighboring countries functioning according the international table. Consequently, it is necessary to consult not only the international table of frequency allocations but also the national table[1] to find out how a particular frequency is being used in a country

The term service in spectrum management carries multiple meanings. It may refer to the engineering concept of radio communication service, which is defined as a service involving the transmission, emission and/or reception of radio waves[2]. Understood as such, radio-communication services are classified according to the nature of the transmitter and/or receivers and the communication features. Examples are the mobile service, the land mobile-satellite service, the broadcasting service, the fixed service etc. In a different context, the term service can be used to describe the functionality that the spectrum usage is intended to provide. For example, television, voice telephony, internet access, etc. Finally, it can be associated to the administrative license required. As so, services can be of exclusive use, collective use, shared use etc. Within the latter definition, there are several types of spectrum licenses which define different kinds of spectrum use.

- Collective use of spectrum: Bands designated for collective use can be utilized without administrative license by anyone. Devices using such bands must comply with a set of standardized technical requirements. Harmful interference among devices is prevented by the fulfillment of the standard and the short ranged nature of transmissions. Collective usages are normally intended to provide short ranged applications that usually don´t exceed a few meters which limits the probability of interference. Examples of applications are in home internet wireless distribution via Wi-Fi, communication of nearby devices via Bluetooth or RFID applications such as payments with smartphones. Longer ranged applications are possible but much less common. An example is the provision of commercial internet access via Wi-Fi technologies.

- Exclusive use of spectrum for commercial purposes. Bands designated for exclusive use can only be utilized for those that have previously obtained an administrative grant. There exists a variety of applications and several types of licenses. Interference among commercial users is prevented by defining in the license the technical conditions under which the radio stations will operate. Frequencies allocated for commercial purposes may be used for the provision of services to the public, like television or wireless broadband access, or be the support of corporate activities whose goal is the production of a commercial good or a service. Examples of supporting activities are wireless transportation of data with radio links or the deployment of private communication networks not publicly available. In some cases, especially in those where applications are intended to provide services to the public, demand for frequencies greatly exceeds supply. In such situations, licenses are normally granted using competitive processes like auctions or public tenders.

- Exclusive use of spectrum for public purposes. These bands are also designated for exclusive use but unlike those described in the previous paragraph, applications are pure public goods, usually can waive any spectrum access or usage fee and are normally provided using public funding. Another differential feature is that bargaining power of public users is lower in case of interference disputes with private applications because of the public nature of the services being provided; high social value and no financial revenue.

- Shared use of spectrum. The shared use of spectrum is an innovative approach to spectrum management where the incumbent users of a spectrum band which is allocated to be used exclusively, share the frequencies with entrants operating under a well-defined set of technical conditions. These conditions are defined in order to avoid harmful interference to existing services. After the introduction of the shared use, both the incumbents and the entrants are able to provide certain quality of service previously defined.

Licenses define the spectrum management model. In practice, the four types presented above coexist in different frequency bands. Some authors have studied these models to ascertain which one is the most efficient. Faulhaber, G.R. (2005) analyzes these models in terms of the capacity of the system to allocate the spectrum to those that value it the most, the minimization of transaction costs, the availability of mechanisms to resolve conflicts of interference and the capacity to adopt new technologies. He concludes that the exclusive use of spectrum complemented by competitive processes like auctions to allocate frequencies is the most efficient model and proposes that even the provision of public services depends on the purchase of exclusive property rights. However, other authors, Cave, M. (2002), ITU (2009) claim that auctions as the only allocation model is not an efficient allocation because the definition of efficiency should include social considerations such as the existence of externalities and pure public goods. Others have suggested collective use as a competing alternative to the exclusive use for the provision of services, Noam (1998), Benkler (2002), Werbach (2004). Collective use of spectrum has very low entry barriers and it has been associated to a boost to innovation but this way has not been explored or has not been successfully enough to be standardized at a large scale. An alternative solution is the exclusive use of spectrum complemented with administered incentive pricing, which consist of charging a fee to spectrum users based on the opportunity cost of using this portion of spectrum, Smith-Nera(1996), Independen et al (2004) and Cave, Doyle and Webb (2007).

3. New approaches to spectrum management

New spectrum management approaches are intended to ease technology migration, reduce transaction costs and achieve an efficient provision of services. Some proposed drastic changes to current management procedures. For example, Faulhaber, G.R. (2005) suggested substituting public allocation and management of spectrum by the definition of exclusive property rights and trust in law enforcement in courts to solve conflicts of interference among parties. The regulator would only be responsible for the definition of property rights including the description of the technical features associated to the use of a frequency band. However, this approach does not take account of the potential under provision of pure public goods or how to include new forms of interference caused by new technologies in the definition of property rights. A list of more gradual and already implemented new approaches to spectrum management is offered below.

- Enabling secondary market transactions like transfers, leasing, and the mutualisation of all or a part of the spectrum license
- Enabling technology migration without prior administrative authorization
- Using auctions as a method of ensuring productive efficiency
- Introducing incentive administered pricing which consist of charging a fee to spectrum users based on the opportunity cost of using the spectrum they have allocated.
- Enabling the shared use of spectrum.
- Defining licenses in terms of acceptable interference to reduce the cost of future transactions of spectrum in the secondary market.

The optimal management approach for each portion of spectrum must be analyzed on a case by case basis. For some bands, the introduction of secondary market transactions may be the most efficient measure but in other cases a better course of action may be to enable technology change in licenses so that the adoption of new technologies is not delayed.

One salient example of a reallocation process carried out using new approaches to spectrum management is the migration from analogue to digital television and the subsequent allocation of the newly available frequencies to mobile broadband. The provision of television services with digital technologies resulted in at least fourfold reduction of the amount of spectrum required to broadcast a television channel with analogue technologies. The new spectrum made available by the migration to digital is usually referred as the digital dividend. The digital dividend was used not only to increase the number of television channels but also to allocate the remaining spectrum to a new service, mobile broadband, via auctions. During the transition period, a repackaging process of the frequencies used by broadcasters was necessary in order to make available a contiguous set of frequencies for the provision of mobile broadband services.

Many studies were carried out on this topic in order to determine how much spectrum of the digital dividend should remain for the provision of television services and how much to be allocated to mobile broadband. Costs benefit analysis was carried out by Analysis Mason et al (2009), Spectrum Value Partners (2008) and Ofcom (2006). All of them focused on the valuation of the private value of the allocation of an additional unit of spectrum to each service. However, none offer a systematic approach to take account of the value of externalities associated to services. The first study lack of the inclusion of externalities in the analysis and the other two consider that the external value of the analyzed services is a tenth of the private value but don´t offer the rationale for such conclusion. However, the external value of services may be a significant part of the spectrum value. For example, many studies associate the provision of broadband to productivity increase and economic growth, Gillet, S. et al (2006), Litan y Rivlin (2001), Goss (2001), López Sánchez et al (2004, 2006) and ESADE (2012). An additional shortcoming of these studies is that they neither include the analysis of the cost of the potential interference problems that may happen during the reallocation process.

4. Methodology to evaluate the efficiency of the reallocation of a portion of spectrum used by incumbent users

We offer a methodology to evaluate total welfare of a reallocation of spectrum. Such methodology is useful to analyze different alternatives and it helps to determine the most efficient allocation of a spectrum band that is currently being underused. New spectrum management approaches normally lead to situations of underutilization of spectrum because ease the adoption of new technologies and reduce intra-service transaction costs making new portions of spectrum available for new allocations.

There are several circumstances leading to the reallocation of services into new available frequencies. One possible scenario is the adoption of a new technology enabling the provision of a service with fewer amount of spectrum. In such cases, the spectrum surplus may be distributed between the current users and entrants with higher social value. The problem here is to determine the most efficient solution taking account of the two relevant dimensions of the problem. One is which new entrants should have access to the new frequencies. The other is how much of the spectrum available should be allocated to the entrant services and kept for the incumbents. The final goal is to allocate the spectrum in such a way that the marginal benefit obtained through the allocation of spectrum to the entrant equals the marginal opportunity cost of the incumbent relinquishing a portion of the spectrum and accepting potential interference problems. The costs of the transition period should also be included in the cost benefit analysis calculation. The marginal value of a frequency unit is the additional cost borne or avoided by an efficient user of spectrum when it renounces or it is enabled to use an additional unit of spectrum.

Another probable scenario is the identification of a band to be shared among several users, the incumbent user and one or more entrants. This scenario is to be considered if the expected benefits of spectrum sharing exceed the associated costs. The use of the same band by two or more different services may require technical and regulatory arrangements to ensure that both, the incumbent and the entrant services are provided with sufficient quality.

The value of each of the alternatives considered in the evaluation of the most efficient allocation depends on the expected demand for services, the current deployment of networks -it is easier to deploy networks in neighboring frequencies- and the potential substitutes of the analyzed band. The substitutability of a spectrum band is a key element in the evaluation of alternatives since in some cases spectrum can be an optional input but in others it may be the only technical and financial available solution to provide a service. Factors such as the land layout, network topology and the technology used to provide the service have an impact on the frequency reuse distance, thus in the existence of alternative inputs substituting the spectrum.

There are some important facts to take account of in the evaluation of alternatives. For example, there are decreasing marginal returns to the allocation of spectrum to services. The marginal benefit that can be obtained through an additional allocation to a service diminishes when the total amount of spectrum allocated increases. Furthermore, the utility of the spectrum is variable and depends on the portion of spectrum being considered. Some portions are in great demand and consequently have higher prices. Finally, another important fact is that the provision of some services requires the use of different portions of spectrum. In such cases, there will be variable substitution margins for each band.

When a service is deprived of a portion of spectrum or it has to share it, there is a cost. The opportunity cost of renouncing to use a frequency band is the cost that a spectrum user has to borne as a result of not being able to use this frequency. The cost of sharing a frequency with an entrant is the expenditure made by the incumbent to maintain the same level and quality of service when the entrant have access to the use of this frequency. Opportunity cost calculations should include capital and operational expenditure considering all the potential substitutes of spectrum.

In some cases it may not be possible to provide similar quality of service or coverage to the same amount of people with a reduced portion of spectrum. In such cases, demand will not be met and there will be a reduction of income that should be included in the calculation of the opportunity cost. Additionally, there may also be externalities not enjoyed that should also be added to the calculations. The alternative with lower cost is the one that should be accounted in the calculations. The following are substitutes of using a frequency band.

- Providing the service using an available frequency that is less demanded.
- Employing other inputs that can replace the use of spectrum. For example, deploying denser networks or building wired systems where it is possible.
- Migrating to a new technology that uses the spectrum more efficiently.

Reallocation of spectrum is conditioned by the size of the frequency band required to provide the entrant service. Consequently, it depends on the availability of a sufficient amount of spectrum to accommodate the new service. In addition to that, the creation of guard bands of unused spectrum between services to avoid interference among them it is usually required. Furthermore, the full-fledged operation of an entrant service for effective competition to exist should be enough to accommodate several service providers. All these requisites need to be contemplated when considering the allocation of spectrum to a new service.

Existing infrastructure and frequency allocations are relevant parameters to determine how beneficial it is the new allocation to a service. The cost of using a new frequency band is heavily reduced when there exists infrastructure using nearby frequency bands. Cost saving stems from the possibility to share the passive infrastructure, e.g. towers, rights of way, and even active infrastructure, e.g. antennas. Furthermore, receiving devices are cheaper to produce because of similarities in the manufacturing process that enable to take advantage of the economies of scale. Consequently, a successful analysis of the most suitable service to be allocated in a newly available spectrum should include the study of the bands that have already been allocated to this service and the infrastructure that has already been built using frequencies close to the new allocation. Conversely, when an incumbent service is forced to renounce to some frequencies that are currently in use, the cost of using less frequencies will be minimized if the service can keep part of the reallocated band. This is because existing infrastructure can be re-used, thus, reducing the impact of the reallocation.

The external value generated by services is an important element to add in the cost benefit analysis of a reallocation of spectrum. Externalities can be associated not only to the public usage of spectrum but also to commercial services. Furthermore, externalities are associated not only to the availability but also to the quality of service. Some spectrum usages, for example those provided by electronic communication networks (ECN), can suffer congestion in areas with high population density. When networks are congested quality of service declines and demand cannot be met resulting in a loss of both external and private value. Congestion can be solved with the allocation of additional spectrum and/or the deployment of new infrastructure because these factors determine the number of stations per unit area. The optimal size of infrastructure/spectrum allocation is that which enable to meet demand making marginal cost of consumption zero under the assumption that meeting demand is profitable for the company providing the service. When it is not financially profitable but still socially desirable coverage obligations should be set in spectrum licenses in order to take advantage of the external value of service provision.

The net benefit of a reallocation scenario depends on the benefits obtained by the entrants, including the private and the external value associated to the use of spectrum, the cost borne by the incumbents, the opportunity cost of not using the available band for other purposes and the cost incurred during the transition from the existing allocation to the new one. We offer a methodology to analyze the efficiency of the reallocation of spectrum to radio-communication services.

The benefit gained from the allocation of N new services into a portion of spectrum that is already allocated to M incumbent users can be derived from expressions 1 and 2.

$$\Pi = \sum_{i=1}^{N} \varepsilon_i + \pi_i \qquad (1)$$

Where

$\Pi \geq 0$ is the total benefit obtained from allocating N new services
N is the number of new services to be allocated in the analyzed portion of spectrum
$\varepsilon_i \geq 0$ is the value of the additional externalities gained from the allocation of spectrum to i
$\pi_i \geq 0$ is the private value gained from the allocation of spectrum to i

The cost borne by M incumbent spectrum users when their portion of spectrum is also allocated to N new services is described by the following expression.

$$\mathrm{I} = \sum_{j=1}^{M} \varepsilon_j + \pi_j + \phi_j + \psi_j \tag{2}$$

Where

I is the total cost of allocating N new services
M is the number of incumbents using the analyzed portion of spectrum
$\varepsilon_j \leq 0$ is the loss of external value from the incumbent j due to the re-allocation
$\pi_j \leq 0$ is the private value missed by an incumbent j in the re-allocation
$\phi_j \leq 0$ is the cost borne by j during the transition period when a new service is allocated
$\psi_j \geq 0$ is the private value added through productive efficiency gained by j when a new service is allocated

Total welfare obtained from a spectrum re-allocation is defined by the following expression

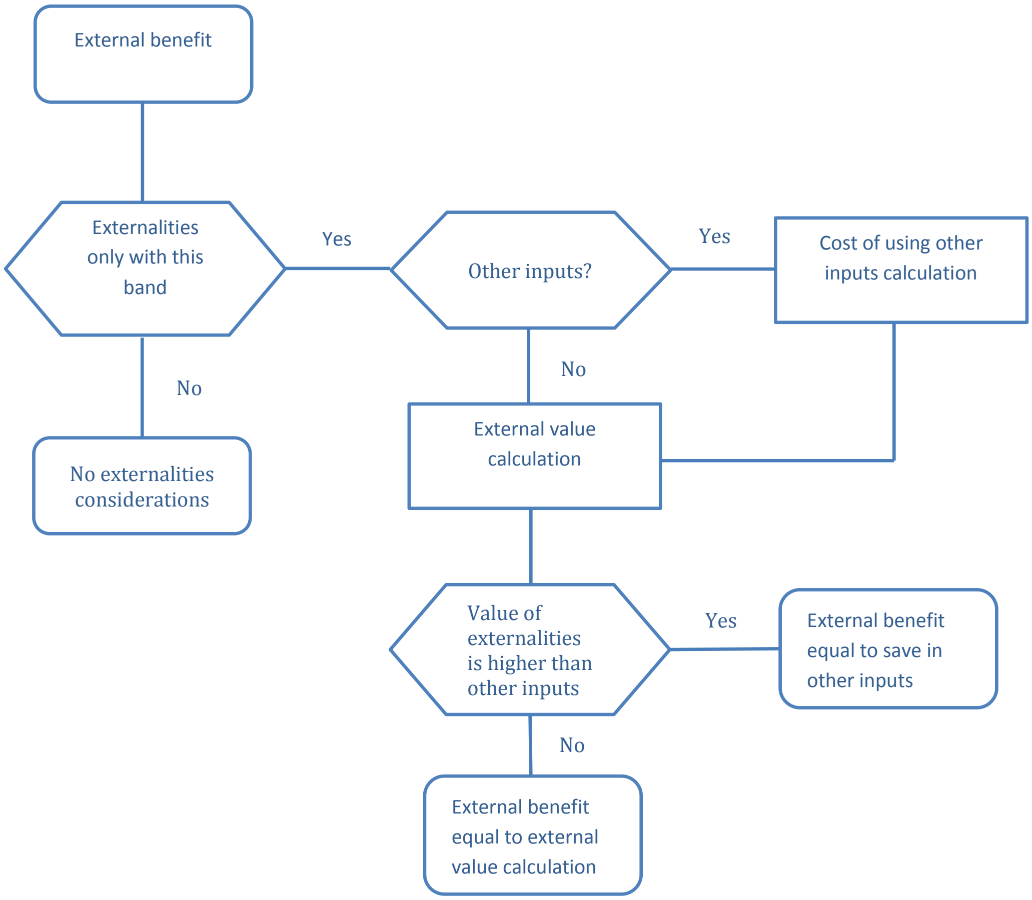
$$TW = \Pi + \mathrm{I} \tag{3}$$

In the following sections we analyze the elements of each of the benefit and costs components and offer flowcharts to help determine the components of the cost benefit analysis.

4.1 Value of the additional externalities gained from the allocation of spectrum to new services

In the calculation of the external value of the reallocation of a frequency band we consider the lowest value among these two; the net positive externality associated to the allocation to a new radio communication service or the cost of building infrastructure for this service to meet demand up to the point where the marginal consumption is zero or demand is completely met.

Figure 1.1: External benefit of spectrum re-allocation



Source: Authors

4.2 Private benefit of reallocating a spectrum band to a new entrant service

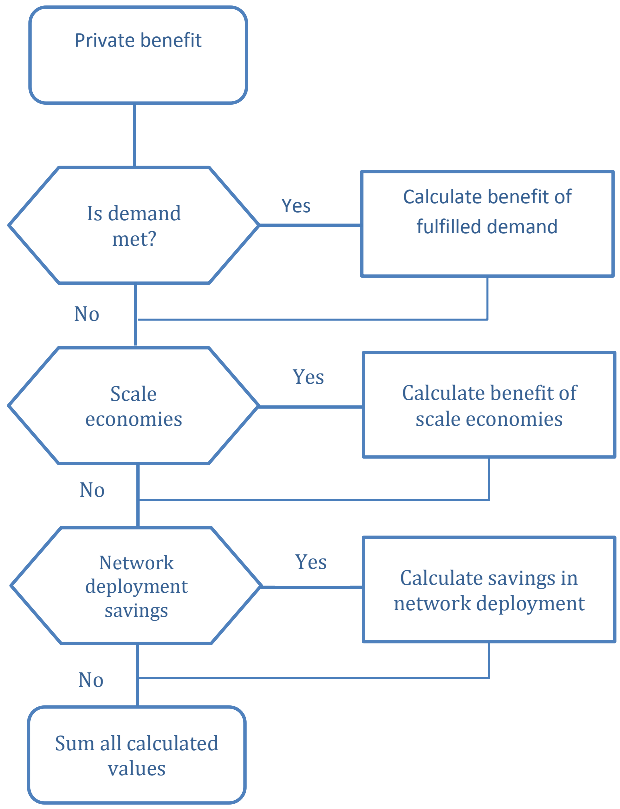
The private benefit of the allocation of spectrum to an entrant service should be calculated for several scenarios considering alternatives for the expected demand and for the existing substitutive inputs of the analyzed band. Producer surplus of service providers should only be included in the calculations in case demand can only be met if the band is allocated. In such cases, it would be added only the additional surplus attributable to the new allocation. If the consumer surplus is not added in the calculation, the net private benefit is the cost saved in the provision of the service stemming from the allocation of the band. Savings might come from capital or operative expenses. Its magnitude depends on the existing infrastructure, the frequencies already allocated to this service, the economies of scale that can be achieved in the manufacture of devices and cost savings gained in service provision.

Economies of scale in the manufacture of the equipment required to use a frequency band can be captured if the band is planned to be allocated in a harmonized way across a large geographical area with high expected number of customers. Cost savings stem from the reduction of the research and development cost per unit of the electronic parts required to use the band and the reduction in the number and complexity of the components embedded in communication devices. If new frequencies are close to those already being used there may be additional cost savings and a reduction of the launching time of new equipment because of the similarity of the new devices with existing ones. Just in time availability of devices is a key element to the successful implementation of the new allocation. Normally, the benefits of the economies of scale are distributed among final users and equipment manufacturers.

Besides the cost reductions in device manufacturing, economies of scale may also apply to service provision. For example, the services provided by ECN can be used by customers visiting a foreign country[3] only if the devices released in their home country share the frequency bands with those in the visited country. Only if such condition is met, service providers can reach roaming agreements to enable service provision abroad. Consumers benefit from the harmonized allocation because they obtain higher utility from the services purchased. The surplus for producers comes from the increased number of consumers demanding the service, country visitors, and from the demand of additional services of local consumers when travelling abroad.

Spectrum allocation save operative and capital expenditure because of the following reasons. It avoids the construction of new sites to meet demand where the network is congested. It can heavily reduce the cost of deploying networks when there exists infrastructure using nearby frequencies because of the possibility to share most of the passive infrastructure and part of the active. Cost savings amount to the differential cost of providing the service with and without the band. Both, the possibility of providing the service with other frequency bands and other inputs different than the spectrum should be considered in the calculations.

Figure 1.2: Private benefit of spectrum re-allocation



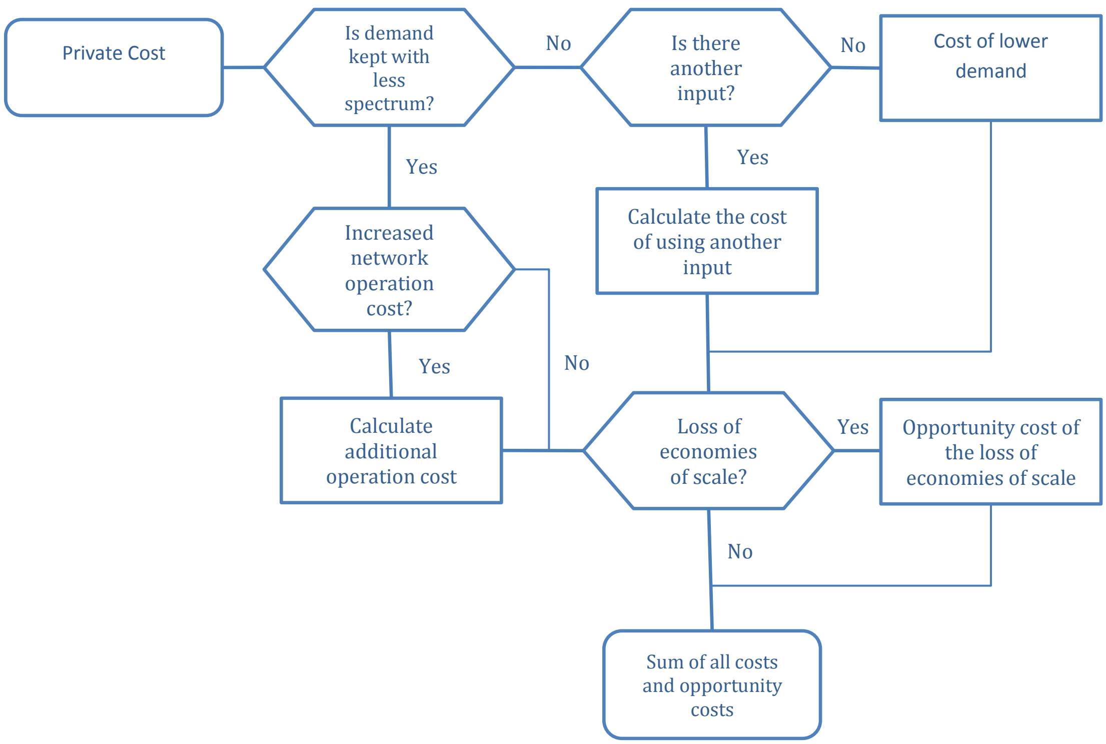
Source: Authors

4.3 Private cost of reallocating a spectrum band to a new entrant service borne by an incumbent service

The private cost of reallocating a spectrum band to a new entrant, is the opportunity cost of not allocating the band to the incumbent or to a different service. Cost calculations should take account of scenarios of the expected demand of the incumbent service and the alternative inputs substituting the analyzed band.

Producer surplus loss should be included in the opportunity cost calculation if the allocation of the band to a new service results in a reduction of the existing supply that cannot be solved by using an alternative input. If it is possible to mitigate the reduced supply with another input, the cost of using this new input should also be included in the calculation. Besides, if the abandonment of the band reduce the current or future expected economies of scale in manufacturing devices or in service provision, these cost also should be accounted.

Figure 1.3: Private cost of spectrum re-allocation

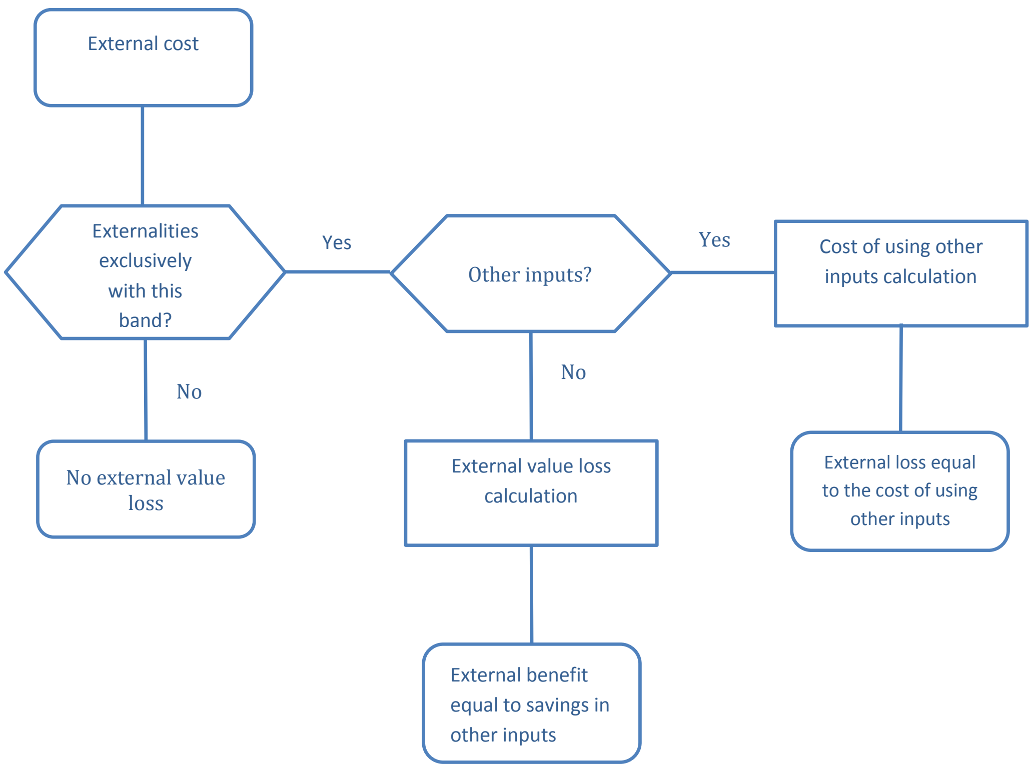


Source: Authors

4.4 External cost of reallocating a spectrum band to a new entrant service borne by an incumbent service

The external cost of reallocating a spectrum band to a new entrant, is the loss of external benefit not enjoyed because of the allocation of the band to a different service. If the allocation of the band to a new service results in a reduction of the existing supply that cannot be solved with the use of an alternative input, the external benefit loss should be included in the opportunity cost calculations. If instead it is possible to mitigate the reduced supply with another input, the cost of using this additional input should be included in the calculations.

Figure 1.4: External cost of spectrum re-allocation



Source: Authors

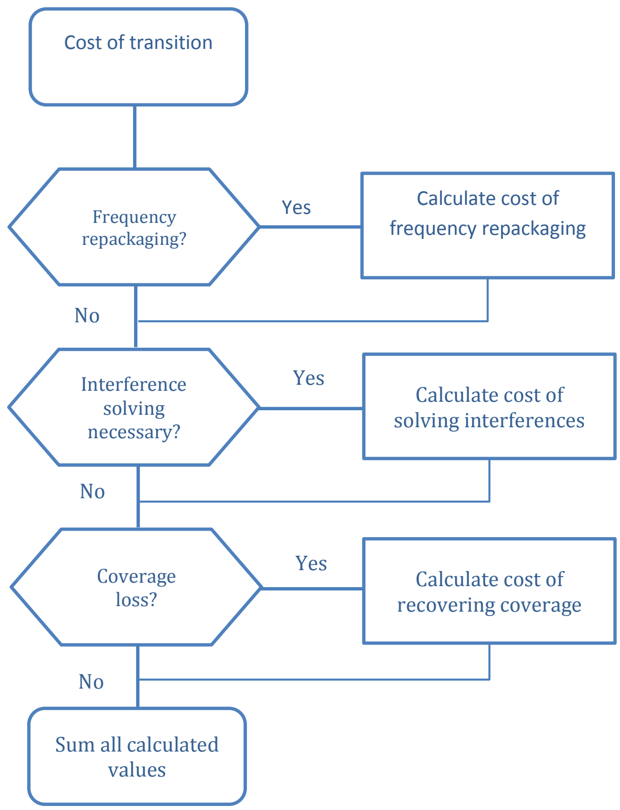
4.5 Cost associated to the transition to a new spectrum allocation

There exist transition costs associated to the reallocation of spectrum to a new entrant service. For the process to be beneficial these costs must be lower than the benefits accrued by the reallocation. Transition cost may be one or several of the following.

- Costs of repackaging frequencies used by the incumbent service to make available a sufficient and adjacent amount of spectrum to be usable by the entrant. Such repackaging activities may imply the adaptation of transmitting stations and/or receiving devices to change the frequency that is currently being used.
- Operating expenses stemming from the provision of services in the new and old frequencies in order to avoid service interruption during the transition period.
- Minimizing harmful interferences that may appear during the deployment of the entrant's network.

- Costs associated to the loss of coverage. These costs that may be solved by using alternative inputs such as other frequency bands or inputs different than the spectrum.

Figure 1.5: Cost of transition to a new spectrum allocation

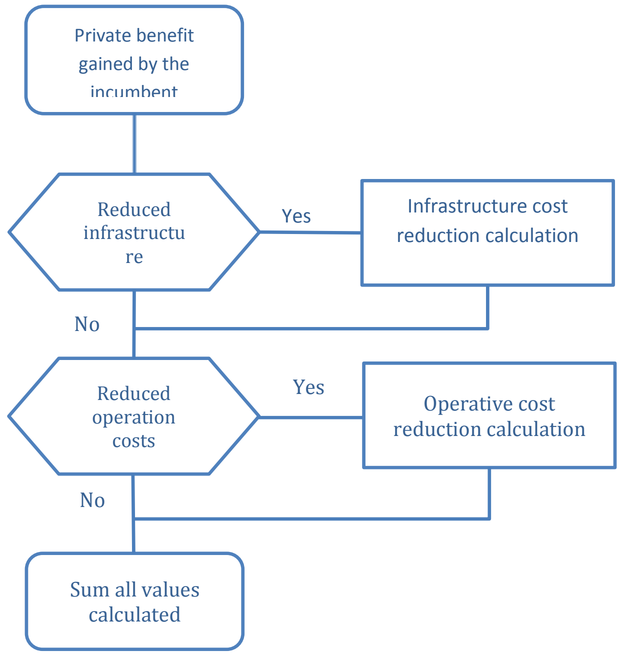


Source: Authors

4.6 Private benefits gained by the incumbent service when reallocating a spectrum band

There may be benefits for the incumbent associated to the reallocation of spectrum to a new entrant service. The reallocation of frequencies might result in the streamlining of frequency usage and service provision reducing operative and capital expenditure. We expect these cost reductions to be several orders of magnitude lower than the opportunity costs incurred by not using the band. Furthermore, if there is a reduction in the number of customers that cannot be solved using another frequency or input, the opportunity cost of the reduced supply should be minored by the reduction on capital and operative expenditure.

Figure 1.6: Private benefit gained by the incumbent with a spectrum re-allocation



Source: Authors

The steps required to apply this methodology are the following

- Identifying a frequency band candidate to be reallocated. Potential candidates are bands where a new technology is either available or already being introduced and bands suitable to be identified for shared use.
- Identifying the geographical area where the analysis will take place.
- Defining the parameters of the cost benefit analysis. Namely the time horizon, the discount factor and other relevant parameters.
- Selecting the entrant services that are suitable candidates to use the new available spectrum. There may be several candidates. All of them should be included in the analysis
- Generating demand scenarios for each service including not only current but also expected demand behavior. Demand for spectrum is driven by the quantity and quality of service. For some services population density is also a driver of spectrum demand.
- Defining a production function for each service under consideration describing how network deployment costs vary with the allocation of frequencies of different portions of the spectrum and with the number of base stations and other inputs.
- Generating the scenarios for potential alternative frequency bands an inputs substituting the allocation of the analyzed band.

- Calculating the value of the additional externalities gained from the allocation of spectrum to new services
- Calculating private benefit of reallocating a spectrum band to a new entrant service
- Calculating the opportunity cost of reallocating a spectrum band to a new entrant service borne by an incumbent service
- Calculating the external cost of reallocating a spectrum band to a new entrant service borne by an incumbent service
- Calculating the cost of transition to a new allocation
- Calculating the private benefits gained by the incumbent service when reallocating a spectrum band

5. Conclusions

The radio-electric spectrum is the essential raw material for the provision of many digital goods and services. These services are changing the economy and society but the benefits of the digital transformation taking place would not be possible without an efficient allocation of spectrum to radio-communication services.

Public management of spectrum allocation is supported by the intrinsic physical features of the spectrum and the nature of some services. Namely, the potential existence of harmful electromagnetic interference among spectrum users, services that are pure public goods, externalities associated to services and economies of scale that can only be leveraged when spectrum allocation is harmonized. However, more efficient approaches and higher market participation in spectrum management has been early and consistently proposed in existing literature.

Spectrum scarcity, especially in certain bands, has led to a consensus among regulators and market agents on the benefits of lighter regulation and higher flexibility in license conditions in order to reduce the cost and speed of network deployments. As a result, many significant developments in spectrum management have taken place. Licenses enabling technology change, the authorization of secondary market transactions and the use of auctions, among other new approaches to spectrum management, have resulted in a more efficient production of services enabling better quality of service, a reduction of the amount of spectrum required, increased coverage and reduced prices.

However, the role of the public sector is still essential in spectrum allocation in spite of the advancements in techniques to increase allocative efficiency, e.g. shared use of spectrum and the definition of licenses in terms of acceptable interference. Technology advancement providing new spectrum resources and improvements in the licensing methods must be complemented with processes of reallocation led by the public sector enabling higher allocative efficiency of resources. These processes should be designed to achieve more transparent transactions and reduced uncertainty for both the incumbent and the entrants to the new portion of available spectrum.

This paper provides a methodology to measure the net benefit of the reallocation of a spectrum band intended to guide regulators and policy makers. After the reallocation, the band will be used by the incumbent and new entrant services. We have identified the following facts.

- The calculation of the benefits and of the opportunity costs of spectrum usage should include the external value associated to the provision of services in cases where externalities can only be obtained using the studied frequency band.

- Demand for spectrum is driven by the quantity and quality of services consumed. For some services the population density is also a driver of spectrum demand.

- The cost of deploying networks to use the spectrum is a key element to determine the private value associated to the use of a frequency band. Deployment costs may be approximated by using production functions. The production function should be composed by several input arguments one of which is the spectrum. There are other inputs such as the number of stations or the use of wired solutions that may be substitutes of a frequency band. If the service is using other frequencies all of them should be included in the analysis.

- The utility of spectrum usage is variable. It depends on the portion of spectrum being considered. Some portions are in great demand and have higher prices. Different portions of spectrum are not perfect substitutes one another, consequently, frequencies have variable substitution margins. The allocation of frequencies with similar features produces decreasing marginal returns. The key goal of spectrum management is to allocate the spectrum so that the marginal benefit obtained through the allocation of spectrum to the entrant equals the marginal opportunity cost of the incumbent relinquishing a portion of spectrum.

- Reallocation of spectrum produces transition costs that should be included in the cost benefit analysis. The cost of transition may include the adaptation of transmitting stations and/or receiving devices in order to change the frequencies currently in use. Operating expenses stem from the provision of services in the new and old frequencies to avoid service interruption during the transition period, costs of minimizing harmful interferences during the deployment of the entrant’s network and costs associated to coverage losses. Sometimes the reallocation may also reduce costs if the provision of the incumbent service and the use of frequencies can be streamlined.

- Economies of scale in the manufacture of the equipment required to use a frequency band can be gained if the band is planned to be allocated in a harmonized way across a large geographical area with high expected number of customers. There may also be economies of scale in service provision.

References

Alden, J. (2012). Exploring the value and economic valuation of spectrum. Report prepared for the ITU. Available at the ITU broadband portal.

Analysys Mason, DotEcon, and Hogan & Hartson (2009). Exploiting the digital Dividend, a European approach. Report for the European Commission.

Baumol, W.J., Robyn, D. (2006). Toward an Evolutionary Regime for Spectrum Governance: Licensing or Unrestricted Entry? Washington, D.C. Brookings Institution Press. AEI-Brookings Joint Centre for Regulatory Studies.

Bazelon, C., McHenry, G. (2013). Spectrum value. *Telecommunications Policy*, Volume 37, Issue 9, October 2013, Pages 737-747, ISSN 0308-5961.

Benkler, Y. (2002). Some Economics of Wireless Communications. *Harvard Journal of Law & Technology*, vol. 16, num. 1, p. 25-83.

Broadband Commission (2014). The State of Broadband 2014: Broadband for all. ISBN 978-92-61-150716.

Broadband Commission (2016). The State of Broadband 2016: Broadband catalyzing sustainable development. ISBN 978-92-61-21771-6.

Cave, M. (2002). Review of radio spectrum management. Report for Department of Trade and Industry and HM Treasury.

Cave, M., Doyle, C. and Webb, W. (2007). Essentials of Modern Spectrum Management. Cambridge University Press. 978-0-521-87669-8.

Cave, M., Pratt, N. (2016). Taking account of service externalities when spectrum is allocated and assigned. *Telecommunications Policy*. Volume 40, Issues 10–11, October 2016, Pages 971–981

Coase, R.H. (1959). The Federal Communication Commission. *The Journal of Law &Economics*, vol 2, p. 1-40.

Coase, R.H. (1960). The Problem of Social Cost. *The Journal of Law & Economics*, vol. 3, num. 1, p. 1-44.

Esade, Penteo (2012). Las Tecnologías de la Sociedad de la Información en la Empresa Española. Informe ESADE-Penteo.

Faulhaber, G.R. (2005). The Question of Spectrum: Technology, Management, and Regime Change. *Journal of Telecommunications and High Technology Law*, Vol. 5. p. 111-158.

Gillet, S. E., Lehr, W.H., Osorio, C.A., Sirbu, M.A. (2006). Measuring broadband's economic impact. Final Report prepared for the U.S. Department of Commerce, Economic Development Administration.

Goss, E. (2001). The Internet´s contribution to U.S. productivity growth. *Business Economics*, vol. 36, num.4, p. 32-42.

Hazlett, T.W. (1998). Spectrum Flash Dance: Eli Noam's Proposal for 'Open Access' to Radio Waves. *Journal of Law & Economics*, vol. 41, num. 2, part 2, p. 805-820.

Hazlett, T.W. (2003). Liberalizing US Spectrum Allocation. *Telecommunications Policy*, vol. 27, num. 7, p. 485-499.

Independen, Aegis Systems and Warwick business school (2004). An economic study to review spectrum pricing. Report prepared for the Radio-communications Agency, UK.

Kwerel, E. and Williams, J. (2002). A Proposal for a Rapid Transition to Market Allocation of Spectrum. Washington, D.C., FCC - Office of Plans and Policy, Working Paper num. 38.

Levin, H. J. (1970). Spectrum Allocation without Market. *The American Economic Review*, 60(2), 209-18.

Litan, R. y Rivlin, A. (2001). Projecting the economic impact of the Internet. *The American Economic Review,* vol. 91, núm. 2, p. 313-31.

López Sánchez et al. (2004) ¿Pueden las Tecnologías de la Información mejorar la Productividad? *Universia Business Review*, primer trimestre, num. 001, p. 82-95.

López Sánchez et al. (2006). Uso de Internet y paradoja de la productividad: el caso de las empresas españolas. *Cuadernos de Economía y Dirección de la Empresa*. Núm. 2, p. 149-174.

Noam, E. (1998). Spectrum Auctions: Yesterday's Heresy, Today's Orthodoxy, Tomorrow's Anachronism. Taking the Next Step to Open Spectrum Access. *The Journal of Law & Economics*, vol. 41, no. 2, part 2, p. 765-790.

Ofcom (2006). Office for Communications. Digital Dividend Review. Mar. 2006.

Prasad, R. (2015). The production function methodology for estimating the value of spectrum. *Telecommunications Policy*, Volume 39, Issue 1, February 2015, Pages 77-88, ISSN 0308-5961

Smith-Nera (1996). Study into the use of spectrum pricing. Report for the Radio-communications Agency, UK.

Spectrum Value Partners (2008). Getting the most out of the digital dividend. Allocating UHF spectrum to maximize the benefits for European society.

International Telecommunication Union (2009). ICT regulation Toolkit. Module 5. Radio Spectrum Management.

Webb, W. (2007). An optimal way to license the radio spectrum. *Regulatory and Economic Policy in Telecommunications*, N2, mar09, p. 100,112.

Werbach, K. (2004). Supercommons: Toward a United Theory of Wireless Communication. *Texas Law Review*, vol. 82, num. 4, p. 863-973.

Notes

[1] For example the international table and national allocations in the United States can be consulted in the following link https://transition.fcc.gov/oet/spectrum/table/fcctable.pdf

[2] As defined by the ITU in the article 1 of Radio Regulations

[3] Or even in the same country if the network provider does not cover all regions of the country

CHAPTER 2: PATH AND SPEED OF SPECTRUM MANAGEMENT REFORM UNDER ASSOCIATED UNCERTAIN COSTS AND BENEFITS

Abstract: *The unsolved question in spectrum management is no longer if a reform is necessary to enable higher market participation but the optimal path and speed of reform with current and expected technological progress. This paper aims to contribute to answer this question offering a model to ascertain the optimal speed of reform and it sheds some light on the policy options available to reduce the uncertainty of the associated costs and benefits of reform.*

Key words: Radio-electric spectrum, efficiency, externalities, technological change.

JEL Classification: O33, D61, D62.

1. Introduction

Since the launching of the first service using radio spectrum in the late 19^{th} century, wireless telegraphy, we have witnessed the emergence of innumerable new services competing for the access to a portion of the limited radio spectrum resources. Examples of services using the spectrum are radio, television, voice and video telephony, internet, national security and defense, whether forecasting, emergency assistance, guidance for navigation of all kinds of means of transport by air, ground and sea, the exploration of the cosmos and many others. Digital services, many of which use the spectrum, have an increasing impact in the economy. For example, they reduce information and coordination costs in the whole economy - activities became cheaper, faster and more convenient-, ease transactions and increase productivity, World Bank (2016).

The allocation of spectrum to services has been planned by governmental agencies practically since the early days of radio communications. Governmental planning has been the response to the complex definition of property rights, high enforcement and transaction costs linked to the existence of harmful electromagnetic interferences, services that are pure public goods, externalities associated to services, and economies of scale when spectrum allocation is harmonized.

In the last decades, improved technology enabling services to share the same portion of spectrum, a better understanding of interference solving and of spectrum markets and especially the extraordinary growth in the demand for spectrum, have put deep pressure to traditional allocation methods. As a result, the reform of spectrum has become a salient policy issue and multiple new approaches to spectrum management have emerged. Examples of spectrum management reforms are the use of auctions to grant licenses, the authorization of secondary market transactions -transfer, leasing, mutualisation of usage rights- , the use of administered incentive pricing, spectrum sharing techniques and the definition of licenses in terms of acceptable interference parameters.

Some remarkably anticipated very early, the need for more efficient approaches to spectrum management, Coase (1959, 1960), Levin (1970). Many more followed these claims supporting the exchange of usage rights in the market and market participation for a better allocation of radio spectrum, Hazlett (1998, 2003), Kwerel and Williams (2002), Baumol and Robyn (2006). Others proposed collective use of spectrum as an alternative solution to scarcity, Noam (1998), Benkler (2002), Werbach (2004). More recently, a number of authors pointed out the importance of incorporating the definition of acceptable interference in spectrum licenses to reduce transaction costs, Webb (2009), ITU (2009), Kwerel and Williams (2010) and Cave and Web (2012).

Spectrum management reform should not be regarded solely as a conflict where we only have two options; either choosing governmental planning or market liberalization. An important component of inefficiency may account for the spectrum allocation process but yet a significant source stems from the intrinsic nature of radio-spectrum. Spectrum physical features together with the attributes of services provided using the spectrum produce market failure. The response should be a combination of a sound regulatory framework, the internalization of technological progress in regulation to overcome market failure, and market participation in spectrum management in order to increase efficiency of spectrum allocation.

Electromagnetic interference, a feature of spectrum usage, is a negative externality that makes the definition of property rights difficult, the respect for these rights hard to enforce and the cost of inter-service and also intra-service spectrum transactions uncertain. Not only the costs but also the benefits of a spectrum management reform are uncertain due to, inter alia, of the existence of hard to measure external value linked to services. García and Valiño (2013), Bazelon and McHenery (2013), Alden (2012), Prasad (2014) and Cave et al (2016) contributed to define methodologies for estimating the value of spectrum and the externalities linked to services. A more general framework to take account of externalities but yet highly applicable to the topic can be found in Cornes and Sandler (1996).

The unsolved question in spectrum management is no longer if a reform enabling higher market participation is necessary or not but the path and speed of reform with current and expected technology evolution, and market and regulatory structure. This paper aims to contribute to answer this question by offering a model to ascertain the optimal speed of spectrum management reform. We first study uncertainty sources to unravel causes and the potential policy responses since uncertainty of benefits and costs of reform is a key element of the porblem. We continue developing the analogy between the case of spectrum reform and the reform of economies in transition commenced by Minervini (2014). We depart from the model in Dewatripont and Roland (1995) adding the idiosyncrasies of spectrum management, namely the fast evolution of digital technologies using the spectrum, and the existence of market failure in the provision of wireless services. We also gained inside from the literature on investment under uncertain benefits e.g. Dixit (1992), Dixit and Pindyck (1994), and the literature about uncertain costs of reform, Pindyck (1993).

This paper is structured as follows section 2 analyzes the different types of reform and the available policy options, section 3 presents sources of uncertainty, section 4 offers a discussion on the optimal path and sequencing of spectrum management reform and section 5 concludes.

2. Types of spectrum management reform

We define a spectrum management reform as the regulatory change intended to enable increased efficiency of spectrum usage so that total welfare is maximized. Reforms can be categorized by the type of efficiency gain that they are intended to produce. Productive efficiency ensure that a service is provided with no more than the spectrum required, allocative efficiency that the assortment of services using the spectrum maximizes social welfare and dynamic efficiency that the compound of services can be changed when a new higher social welfare allocation emerges.

Reforms take account of regulatory and technology innovation. An innovative regulatory change may consist of a new approach for defining property rights that reduces inter service transaction costs and/or the costs to enforce respect to spectrum usage rights, innovative ways to enable intra service transactions or the internalization of technology advance into spectrum usage rights. Specific policy recommendations exemplifying spectrum management reform types can be found in García-Zaballos and Foditsch (2015).

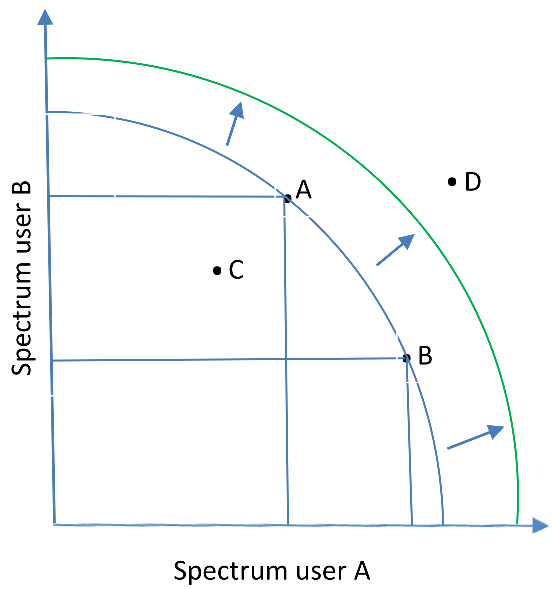
We define two types of technological progress of spectrum usage. A technology advancement improving the delivery of services happens when a new technology allows either using less amount of spectrum to provide a service, or it improves service features and capacity or both at the same time. For example, the second generation of mobile telephony allowed peak speed data rates of 9.6 kbps whereas fourth generation, the so called LTE, allows peak data rates of 1 Gbit/s. Another example is the migration from analogue to digital television. Digital television technologies can provide a minimum of four times more television channels than analogue using the same amount of spectrum. However, new technology such as this, does not improve the allocation of spectrum to services.

A technology advancement improving the allocation of spectrum arises when a new technology enables for a better allocation of spectrum. Better allocations may be obtained from reducing inter service and intra service transaction costs, from allowing the usage of the same portion of spectrum by several services or from reducing the costs to enforce respect to spectrum usage rights. Carrying out spectrum inventories that identify spectrum supply and demand may help to figure out the expected return of a reform. Next section identifies the different types of spectrum management reform and some policy options to implement them.

2.1 Reforms achieving greater productive efficiency

Productive efficiency of a given allocation of spectrum to services is achieved when a firm delivering a service cannot employ an additional unit of spectrum to increase production without reducing the production of another firm delivering the same service. When this condition is met the service is delivered at the lowest average total cost, and it uses the minimum possible amount of spectrum. In other words, the use of a portion of spectrum is productively efficient if service delivery is at the production possibility frontier. Figure 2.1 shows two firms A and B using a certain amount of spectrum to provide the same service. Points A, B are productively efficient, point C is inefficient and point D is not possible with current technology. The figure also depicts the effect of applying a new technology to the production of a service. A new technology improving productive efficiency would move right the productive possibility frontier curve.

Figure 2.7 Production possibility frontier of spectrum use with technology advancement



Source: Authors

Auctions of spectrum rights of use and the authorization of secondary market transactions are spectrum reforms intended to grant spectrum rights to those that value the spectrum the most and can deliver services using the minimum possible amount of spectrum. In figure 2.1, point C is not productively efficient. Auctions and secondary market transactions may move C closer to the production possibility frontier.

One way of achieving higher productive efficiency of a new allocation of spectrum is to auction the rights of use. If licenses have already been granted, the authorization of secondary market transactions may help to improve the initial distribution of spectrum to firms. Examples of secondary market transactions are the possibility to transfer or leasing the spectrum rights of use to a different user providing the same service or the mutualisation of spectrum rights. Mutualisation consists of sharing all or part of the spectrum license among two or more firms providing the same service in the same or in different spectrum bands.

As we can see in figure 2.1, the production possibility frontier curve shifts right if a new technology enables to use less spectrum or it improves service features. In this case, productive efficiency gains can be achieved immediately if spectrum licenses allow spectrum users to change technology without prior administrative authorization. For efficiency gains to be obtained technology change must not cause harmful interference to spectrum users in the same or adjacent bands. An example of an authorization of technology change is the so called spectrum re-farming authorizing the migration of spectrum use to more advanced generation of mobile technologies. If technology change is not allowed and there is technological change, points A and B in figure 2.1 that before the new technology were efficient would be situated below the new production possibility frontier.

2.2 Reforms achieving greater allocative efficiency

Allocative efficiency is achieved when the distribution of services into spectrum portions is optimal so that social welfare is maximized. Optimality includes not only the private value of spectrum, - the willingness to pay of the firm using the spectrum -, but also its social value -externalities linked to the delivery of services-. Allocative efficiency is achieved when the marginal benefit of spectrum use equals the marginal cost of service delivery for all applications using the spectrum.

Achievement of productive efficiency does not preclude the attainment of allocative efficiency. Let's consider a service A using the spectrum with lower value for society than service B. Service A enjoys the allocation of a higher amount of spectrum than B. The introduction of a productively efficient reform in both, A and B, may increase total welfare less than changing the allocation of spectrum in order to grant more spectrum to the service A with higher social value. Table 2.1 shows a spectrum reform increasing productive efficiency in A and B. Table 2.2 represents a reform changing the allocation of spectrum to increase social welfare. As it can be observed total welfare is higher in reform 2 than in reform 1. It is important to note that for the calculations in table 2.2 we have considered diminishing returns to the amount of spectrum allocated to a service. This is because social value per unit of spectrum allocated to a service decreases when the amount of spectrum allocated to this service increases.

Table 2.1: Productive efficiency reform

Service	Units of spectrum allocated	Social value per unit of spectrum	Productive efficiency (Before reform)	Productive efficiency (After reform)	Total Welfare
A	10	1	75%	90%	9
B	2	5	75%	90%	9
A+B	12	1.664 (Avg)	75% (Avg)	90% (Avg)	18

Source: Authors

Table 2.2: Change of allocation reform

Service	Units of spectrum allocated	Social value per unit of spectrum	Productive efficiency	Total Welfare
A	6	2	75%	9
B	6	4	75%	18
A+B	12	3 (Avg)	75% (Avg)	27

Source: Authors

Examples of reforms increasing allocative efficiency are the use of administered incentive pricing, incentive auctions and spectrum sharing licenses. Administered incentive pricing [1] (AIP) consists of charging a fee for the use of spectrum which is related to the opportunity cost of using this portion of spectrum, Smith-Nera (1996) and Cave, Doyle and Webb (2007). Different portions have distinct values, so using AIP requires calculating the opportunity cost of using each frequency band. The purpose of AIP is making license owners to return the spectrum if they are using a band inefficiently. The fee should take account of the costs borne by another services not using the band. Costs may come from using the service with means different to spectrum, or with a less valuable spectrum band. AIP can be a good measure to increase allocative efficiency when it is possible to calculate spectrum value, both private and social. However, a state agency may lack the monetary measure of costs and profits of spectrum use because it does not have all the relevant information regarding consumer preferences and spectrum users' willingness to pay.

Incentive auctions[2] consists of two related auctions, a forward auction where it is determined a firm's willingness to pay to access the spectrum and a reverse auction which establish the price for which the firm using the spectrum is willing to sell it. Prior to the auction, it is necessary to design a frequency repacking process analyzing the potential transaction costs of spectrum reallocation in order to reduce the uncertainty of the process. Incentive auctions enable the change of use of spectrum, thus increasing not only allocative efficiency but also productive efficiency. The allocation change is valued by the market instead of administratively determined, so reducing the possibility to cross-subsidize one of the services involved in the transaction. An example of an incentive auction design can be found in FCC (2012).

The term spectrum sharing encompasses a set of different spectrum management concepts. It may refer to the collective use of spectrum, the dynamic spectrum management, or the underlay spectrum management. Definitions can be found in RSPG (2011). The collective use of spectrum is the use of spectrum without license under a set of well-defined set of technical conditions. Spectrum user can access the spectrum provided that they use devices in compliance with the conditions previously defined. Collective use does not increase allocative efficiency since it can be used neither to migrate to a more efficient allocation of spectrum nor to allow the delivery of several services in the same portion of spectrum. Collective use is a way of delivering services that doesn't require the grant of rights of use. It is normally employed for allocating spectrum to short-range devices such as Wi-Fi or RFID applications.

Spectrum underlay also known as spread spectrum technologies consist of the emission of very low spectral power density signals that can coexist with other spectrum uses in the same frequency, with the effect of slightly increasing the noise floor -electromagnetic interference- to incumbent spectrum users. Spectrum underlay is a way to increase allocative efficiency in spectrum management since it allows the use of the same portion of spectrum by different services and it also enables future changes of use.

Dynamic spectrum management includes both, regulatory instruments and technical approaches. It is intended to allow using a portion of spectrum by an entrant and an incumbent spectrum user taking advantage of the fact that the incumbent is not using a frequency in a particular geographical area or time slot. Examples of technical approaches are cognitive radio, sensing of frequencies being used, geo-location beacons or databases of spectrum usage. A regulatory instrument enabling dynamic use of spectrum is the definition of shared access licenses. Such licenses delimit usage conditions of a particular portion of spectrum being used by an incumbent for a defined time period or area. Table 2.3 shows total welfare after introducing both a new allocation and a spectrum sharing reform.

Table 2.3: Spectrum sharing reform

Service	Units of spectrum allocated	Social value per unit of spectrum	Productive efficiency	Total Welfare
A	6	2	75%	9
B	6	4	75%	18
C	2	2	75%	3
A+B+C	12	3 (Avg)	75% (Avg)	30

Source: Authors

2.3 Reforms achieving greater dynamic efficiency

Dynamic efficiency is achieved when it is possible to change the allocation of spectrum in response to either the emergence of new services with high social value or when there is a change in the value generated by services that turns inefficient current allocation. For example, the increasing number of applications offered through mobile broadband increase the social value of this service over time.

Spectrum reforms intended to increase dynamic efficiency also increment allocative efficiency. The significant difference between reforms increasing allocative efficiency and reforms enhancing both is that the latter can be used not only in a particular case but also in future changes of allocation. The line delimitating whether a reform increases only allocative or also dynamic efficiency is sometimes blurred. For example, the design and processes of an incentive auction may be re-used in the future if the service demanding to access the spectrum share technical features with the services for which the incentive auction was designed, but in other situations the incentive auction should be redesigned.

Examples of reforms dynamically efficient are those enabling future changes of spectrum allocations. Two examples are the authorization of spectrum underlay technologies, and the definition of licenses in terms of acceptable interference. Spectrum underlay technologies are able to use the spectrum that is already in use by other service without interfering with the incumbent user. Future changes of the underlay spectrum use are possible.

The definition of spectrum licenses in terms of the maximum acceptable interference, that devices are willing to accept, facilitates inter-service spectrum transactions, thus the dynamic allocation of spectrum without administrative intervention. Let's define an interference limit that the devices using a frequency band are willing to accept, for example -110 dBm per MHz[3]. Transaction costs of a change of spectrum allocation would be greatly reduced if an entrant service is able to comply with this limit and it is willing to pay for the access to the portion of the spectrum for which limits of acceptable interference has been defined. The lowered transaction costs stems from the reduced risk of interference. Even if a frequency repackaging process is necessary in the reallocation process, uncertainty and transaction cost will have been greatly reduced. An increasing number of authors have claimed for the importance of incorporating the definition of acceptable interference in spectrum licenses, Webb (2009), Kwerel and Williams (2010) and Cave and Web (2012).

2.4 Mixed reforms

Spectrum reforms rarely have economic sense implemented individually. Reforms usually come in packages that need to be carried out sequentially. Table 2.4 shows total welfare of a productively efficient reform followed by a reform improving allocative efficiency. The initial allocation is described in table 2.1.

Table 2.4: Mixed reform

Service	Units of spectrum allocated after reform	Social value per unit of spectrum	Productive efficiency (Before reform)	Productive efficiency (After reform)	Total Welfare
A	6	2	75%	90%	10.8
B	6	4	75%	90%	21.6
A+B	12	3 (Avg)	75% (Avg)	90% (Avg)	32.4

Source: Authors

Normally it is not possible to make a choice on the order of spectrum reforms. However, it is possible to choose the starting time and speed of the first and subsequent changes. To exemplify how a mixed reform is carried out, let's consider technology advancement allowing the provision of a service using less amount of spectrum. The result of reform implementation is a new portion of available spectrum that can be used either to increase the amount produced of the existing service or to allocate a new service with expected higher social value. If the latter option is carried out then a second reform is required to accommodate the entrant service to the available spectrum.

A real example of such situation is the transition from analogue to digital television in the UHF band, the so called digital dividend. Transition resulted in the availability of new frequencies in a band highly demanded. After the transition part of the frequencies of the digital dividend were allocated to mobile broadband. The new allocation required a series of complex technical studies in order to reduce interference problems between mobile broadband and digital television services.

3. Uncertainty associated to spectrum management reform

Spectrum management reforms intended to increase productive efficiency are usually less prone to uncertain results. Consequently, once regulatory mechanisms are created, transactions can be carried out by market agents provided that there is sufficient supervision on effective competition[4]. Examples are transfers, leases, or mutualisation of spectrum rights among firms providing the same service.

The incorporation of technology change in licenses, the so-called frequency re-farming, enables market agents to efficiently carry out inter-band technology change since interference problems that may arise would be solved as described in the Coase theorem. The theorem conditions, well defined property rights and low transaction costs are usually met in spectrum re-farming type of transactions. However, market agent behavior might not always be efficient. If the adoption of a new technology results in increased competition, market agents might not have the incentive to embrace it in order to keep the privileges of reduced competition. In such cases, a regulatory action would be required. Lack of incentives to adopt new technology may explain why in most countries radio FM services are still being provided using analogue technologies while much more efficient digital technologies are available[5].

Reforms intended to change the allocation of spectrum and mixed reforms where an allocation change is included are usually expected to produce the highest benefit if carried out in the right moment. However, such reforms have not only uncertain returns, especially uncertain external benefit fulfillment, but also uncertain costs. Lack of information about the cost of reforms is an important reason for the market to refrain to carry out transactions. Regulatory mechanisms such as the design of incentive auctions or the definition of licenses in terms of maximum acceptable interference might help market agents to increase efficiency of changes of spectrum allocation. The following sections analyze uncertainty of spectrum management reform both in the return and the cost of reform.

3.1 Uncertain return of a spectrum management reform

The return of a spectrum management reform is uncertain because of the existence of externalities, services that are pure public goods and economies of scale when the use of spectrum is harmonized. The following sections analyze each of the causes of uncertainty.

3.1.1 Externalities

The analysis of the social efficiency of a spectrum reform should include the reform effects on economic growth, equal opportunities for marginal groups in society and the promotion of effective competition; see Cave (2002), ITU (2009).

Positive externalities appear as a result of the increased number of possibilities for producing goods and services, easier access to information, reduced transaction costs and new chances for consuming new products and services. If these externalities are not included in spectrum management decisions some services would be provided at a below-the-optimum amount. For example broadband, a service that can be provided using the spectrum, has been associated to GDP growth and productivity increase, Gillet et al (2006), Litan and Rivlin (2001), Goss (2001), López Sánchez et al (2004, 2006).

Production below the optimum may happen for a variety of reasons. One is the economies of density associated to network deployment. Rural areas coverage may be socially desirable but privately unprofitable. Urban-rural cross subsidy or public funding are socially efficient in these cases to achieve the optimal social production. Another one is sunk costs[6], subsidies and/or public private partnerships may help mitigate the problem.

The expected value of a reform may be biased by miscalculating the external value. The social value of spectrum includes what spectrum users are willing to pay plus the externalities associated with the services using the spectrum. Whereas making an educated guess about the private value may be possible, the value of the spillover effects is much more uncertain and difficult to assess.

Private value may be calculated with some precision by carefully studying the result of previous auctions, applying the cash flow method[7], measuring the difference in the total cost of ownership of deploying a network without a spectrum band, Frias et al (2016), or even using production functions see R. Prasad (2014). Sometimes, the calculation of the external value may not be necessary if the loss of spectrum resulting from a reform can be compensated by non-spectrum inputs, e.g. a higher number of radio stations, see J.M. Garcia and A. Valiño (2013), M. Cave (2016). In this case, the calculation of externalities may be substituted by the calculation of the cost of providing the service by other means. In other cases, such substitution is either technically unfeasible or financially impossible since the cost of alternative inputs is unbearable. Then, there is no alternative to calculating the cost of externality loss in order to assess the expected return of the reform.

3.1.2 Pure public goods

Some spectrum services provide goods non-excludable and non-rivalrous. Examples are national security and defense, emergency assistance, scientific research, whether forecasting or climate change monitoring. These services are not delivered by the market and are usually financed with public budget for everyone's benefit. How much spectrum has to be allocated to those services is uncertain since the calculation of the benefits they produce is complex.

Spectrum management reform based on the increased use of spectrum sharing has been proposed in the US and Europe as a means to increase efficiency, especially in bands currently used to provide pure public goods. However, there must be mechanisms to maintain the optimal production of these public goods ensuring both an optimal amount of spectrum and a service provision free from harmful interference.

For example, a US Government Report, PCAST (2012) found that clearing and reallocation of Federal spectrum[8] -the spectrum used to deliver pure public goods- would be expensive and lengthy and reallocation a not sustainable basis for spectrum policy. Instead, it is proposed to use spectrum sharing as a means to increase efficiency under a licensed regime. In Europe, a Communication from the European Commission to the Parliament, EC (2012) defined the conditions to promote the shared use of spectrum.

The definition of clear and effective sharing rules may help to formulate better defined property rights and reduced transaction costs, thus satisfying Coasian schemes. However, the expected result of the Coase theorem, the achievement of an optimal agreement mutually advantageous between spectrum users, may work only for spectrum sharing between commercial uses. In the case of social production, even if public goods have higher social value, the public service provider will not be able to pay for interference solution when it happens. Thus, when the spectrum is shared between services producing pure public goods and commercial goods there must be mechanisms to enforce property rights and avoiding harmful interference.

3.1.3 Economies of scale

The allocation of the same portion of spectrum in a broad geographical area, known as spectrum harmonization, produces important economies of scale which reduces the cost of manufacturing transmitting and receiving devices. Savings stem from the distribution among a larger number of users of the cost of designing the electronics needed for using a frequency band. It also arises from the diminished technical complexity of the device resulting from the reduced amount of frequencies it has to operate with. The latter effect is especially important in services that are supplied using several frequency bands such as mobile telephony.

For some services, there are additional benefits of harmonization. For example, the allocation of the same spectrum band to mobile networks in several countries enables to continue using the same mobile terminal in different countries via “roaming”[9] agreements. In such case, not only devices cost less but also service utility is enhanced. Harmonization of a portion of spectrum may lead to the harmonization of the adjacent portions. The use of an adjacent spectrum band may exponentially reduce the cost of deployment of a new network because it enables re-using infrastructure of existing sites[10] and it creates economies of scale in network deployment.

Regulators and market agents involved should act concertedly to make harmonization of certain bands possible in order to achieve the benefits of the economies of scale. Worldwide harmonization is desirable in certain occasions. The calculation of the return of a spectrum management reform is uncertain since it may depend on the agent coordination success.

3.2 Uncertain cost of spectrum management reform

Changes in the allocation of spectrum may create harmful interference, fragmentation and costs associated to the reallocation of existing services. A high level of spectrum exploitation is associated with higher adaptation costs of existing devices, higher cost for reallocating spectrum, and compensations to spectrum users with licenses still in force.

3.2.1 Uncertain interference levels

Changes in allocation of spectrum may create harmful interference to existing services stemming from the different emission parameters, density of stations and typical power of the new services.

3.2.1.1 Different emission parameters

The relevant parameters involved in the characterization of harmful interference[11] are the design of transmitters and receivers[12] and the emission parameters and techniques[13]. Main emission parameters are frequency, power, location and direction. The emission techniques depend on technology[14]. These techniques are normally intended to increase coverage or service capacity but may also have an impact on the existence of or shielding to harmful interference. Transmitters are designed to produce a certain limited amount of out of band emission[15] without interfering neighboring spectrum users and receivers are able to accept certain amount and still performing well. Consequently, the existence of harmful interference depends on how sensitive to harmful interference are the devices of the incumbent services to the emissions of the entrant and vice versa.

Compatibility between stations providing a service and frequency adjacent services is achieved through the setup of limits on emission parameters that are set forth in the spectrum license by a Governmental agency. Devices and emission techniques are designed to on the one hand maximize capacity and on the other hand to avoid harmful interference among radio stations of the same and adjacent services. If there is a change in spectrum allocation, the existing emission parameters and /or devices may interfere with or be interfered by new devices using the adjacent band since they were not initially designed to deal with the new sources of interference. The cost to cope with interference is uncertain not only from the point of view of the engineering process, complex compatibility studies are required, but also from the point of view of the coordination of all involved agents.

A spectrum allocation approach with reduced interference problems is the use of underlay spectrum sharing techniques. Under such approach, new devices are able to co-exists with incumbent users in the same frequency band under a set of well-defined conditions, but this kind of solutions are limited to low power applications. Dynamic sharing is an additional approach that may be used when incumbent services are not using the spectrum in a particular geographical area or time slot.

A more global solution is the reform of spectrum licenses in order to include the maximum interference that devices should be able to tolerate instead of defining their emission parameters. Doing so, would enable future changes in the allocation of spectrum that would be much less complex and uncertain.

3.2.1.2 Density of stations, the cumulative effect

The greater the density of radio-stations in an area the higher it will be the probability of harmful interference due to the accumulation of interfering signals of multiple nearby stations. Webb (2009) maintains that when there is a change in spectrum allocation the cumulative interference effect may have a higher impact on the entrants than on the incumbent service. Networks using the initial allocation of spectrum are usually deployed in parallel, thus, the density of stations is similar and the cumulative interference problems are solved dynamically during network deployment. However, when there is a new allocation of spectrum, the network of the entrant service is not yet deployed and it is more exposed to the potential cumulative inference of a much denser network in an adjacent band. The cumulative effect increases the interference problems suffered by the entrants.

A well-designed spectrum license should incorporate a set of parameters including not only the maximum acceptable interference from a station but also taking account of the cumulative effect. Such parameter definition complicates license design but diminish future transaction costs.

3.2.1.3 Differences in the typical power of stations

Problems such as coverage holes[16] or receiver's blocking[17] may appear if there is a significant difference between the typical power of stations of the incumbent service and the stations of a new spectrum allocation. When the typical power of stations is similar, the location of both the entrant and the incumbent services' stations tend to be similar and such problems are much less important.

A solution to the problem may be to create different frequency bands bearing in mind the typical power of stations providing the service. These frequency bands would be separated by guard bands to mitigate potential interferences among services. Within these bands, changes in the allocation of spectrum to services with similar features would be much easy to be carried out. A potential optimal organization of such frequency groups would be the creation of bands for unidirectional high power networks - broadcasting-; for unidirectional low power- mobile multimedia-; and low power bidirectional networks- mobile broadband.

3.2.2 Fragmentation of spectrum allocation

Reallocation of spectrum might create gaps and fragmentation of spectrum use. The minimum number of spectrum units used by each service is usually different. Consequently, reallocation may result in unused portions of spectrum. For example, let's consider service A using multiples of eight units of spectrum to provide a service and service B using multiples of five. If service A is willing to relinquish eight units in exchange for compensation and service B is willing to pay the compensation, three units would be left unused until a new service is willing and it is able to use them. If service B would require 10 units, then the result is either the impossibility to make the transaction or the creation of a band of six units of unused spectrum, provided that B is willing to relinquish 16 units.

Furthermore, the provision of services on adjacent frequency bands usually requires the creation of a guard band[18] whose size may be one or several units of spectrum in order to avoid harmful interference among services.

3.2.3 Uncertain costs associated to the existing licenses and devices

A higher level of usage of a frequency band by incumbents is associated with higher adaptation costs of existing devices, higher cost for reallocating the spectrum, and compensations to spectrum users with licenses still in force when there is a reallocation process.

3.2.3.1 Compensation costs

Reforms changing spectrum allocation may pose important compensation costs. In some cases licenses have not expired and current licensees have the right to receive compensation for loss of earnings before a spectrum band is made available to a new service. When licenses have different expiration periods it may be more difficult to make the frequencies available, Minervini and Piacentino (2007).

3.2.3.2 Reallocation costs during the transition period

Reforms making available a portion of spectrum tor a new service involve financial costs beyond the compensations for loss of earnings. Such reforms require adaptations like changing the transmission frequency, or adapting receivers so that they are no longer able to receive the frequencies allocated to the new entrant service. Receiver adaptation can be done by placing filters at the input of the receiver. Additionally, in some cases it may be necessary to simulcast[19] new and old frequencies during a period which imply higher network operation costs for the incumbent during the transition period.

3.2.3.3 Adaptation of existing devices

The sensitivity to the adjacent channel interference is an important parameter of a receiver. The more resilient it is the device the more expensive it becomes the electronics used to mitigate interference. Manufacturers design devices to work in a certain predefined environment. If a new allocation is made in an adjacent band, receiver features may not be sufficient to avoid harmful interference produced by the entrant services. There exist mitigation techniques such as placing filters at the input of the receiver to solve the problem. However, the total cost of the adaptation is uncertain because depends on the features of the stock of receivers, which usually is an unknown variable. Once again, the definition of licenses in terms of acceptable interference would help to manufacture more resilient devices, consequently, lowering transaction costs of changes of spectrum allocation.

4. A model to analyze the path and speed of spectrum management reform

We bring the literature of political economy of reform to the topic of spectrum management. We continue developing the analogies between the case of spectrum reform and the reform of economies in transition commenced by Minervini (2014). We depart from the model in Dewatripont and Roland (1995) adding the idiosyncrasies of spectrum management. First, we include the evolution of technology as an endogenous variable and then we allow for different discount factors for each reform depending on technology evolution. We have changed some of the notations, explained the meaning of some variables from the view of spectrum reform and restricted the value that some parameters can take but assumptions and definitions remain similar to those used in Dewatripont and Roland (1995).

The most significant difference between the reform of spectrum management and of economies in transition is the pace of technological progress. Whereas in the latter technology is 'sticky' and it doesn´t change much during the timeframe of reform implementation, in spectrum reform it may change during the analyzed timeframe. Technology is a salient factor to determine the strategy of a spectrum reform. It determines whether to implement gradual or big bang type of reforms. Consequently, in our model technology is an endogenous explanatory variable. Furthermore, the realization of states of nature, the exogenous variables, changes when reforms are carried out with new technology since technology acts as a multiplier of some states of nature.

A spectrum management reform is a set of regulatory changes intended to increase productive, allocative or dynamic efficiency. In most cases, the order of reform implementation cannot be chosen but reformers can decide when to start and the speed. The starting moment is important because not only costs but also the benefits of reform depend on the technology used when the reform is carried out.

Big bang reform packages are those introducing reforms quickly and simultaneously. Gradual reform packages are those implemented in such a way that it can be decided whether to proceed with the next package or postpone it. One reason to postpone a reform is waiting for the arrival of a new expected technology.

Reforms are regulatory changes intended to enable increasing efficiency either productive or allocative or dynamic. Examples were described in section 2. Reforms are carried out sequentially if they are complementary. Although we simplify the model considering only two sequential reforms we do not loss generality since it can be understood as complementary packages of reforms that can be carried out in two sequential moments.

Reform packages can be designed to increase any of the types of efficiency. There may be two productively efficient reforms in different portions of spectrum or geographical locations. For example, the use of auctions in a portion of the spectrum followed by the use of auctions in the rest of portions being used by this service. It can also be mixed reforms packages where it is necessary first a productively efficient reform and then a subsequent reform improving allocative efficiency. An example is the introduction of a new technology that free up spectrum that can be used by an entrant. Finally the definition of spectrum licenses in terms of the maximum acceptable interference may be carried out for a spectrum band and hereinafter in other bands.

The outcome generated by a reform depends on a set of possible states of nature with N_i elements. These elements include the uncertain factors described in section 3.2, interference levels, fragmentation of spectrum, costs associated to the existing licenses and devices and also when applicable what was described in section 3.1, economies of scale, externalities and pure public goods. Other aspects such as political constraints, investment behavior are also included in the composition of the states of nature but for those, technology is not a multiplier.

We consider an infinite time horizon. Reforms are carried out with a discount factor $\delta < 1$. There are two reforms $i = 1{,}2$ to be carried out in the period analyzed. The outcome of reform i depends on the realizations of the states of nature. The realization of the k th element of a reform is $s_{ik} \in \{s_{i1}, s_{i2} \ldots, s_{iN_i}\}$. Therefore, the realization of the k th element of reform 1 is s_{1k} and the realization of the m th element of reform 2 is s_{2m} . The corresponding outcome for these realizations is $o(s_{1k}, s_{2m}, t) = o(s_{1k}, s_{2m})$. We consider reform outcome to be time invariant. The present value of $o(s_{1k}, s_{2m})$ at the moment when the first reform is realized is denoted by $O(s_{1k}, s_{2m}) = o(s_{1k}, s_{2m})/(1 - \delta)$. Technological advancement is denoted by $\sigma > 1$.

The outcome of a reform is unknown before it is implemented, for this reason we work with ex ante expected results. We denote expected outcomes depending on the strategy of reform, big bang or gradual, and the knowledge about the payoff of the first reform once it is realized. $E_{k,m}O(s_{1k}, s_{2m}) \equiv O_{k,m}$ is the expected outcome of a big bang reform before any of them have been realized. $E_m O(s_{1k}, s_{2m}) \equiv O_m$ is the expected outcome of a big bang strategy given that the first reform has already been conducted. $E_{k,m}H(s_{1k}) \equiv H_k$ is the expected outcome of a gradual type reform once the first has been realized. $O_m > H_k$ because of the required complementarity of reforms. However, in gradual reforms, from the implementation of H_k one can learn about the final total payoff of reform and postpone second reform, whereas in big bang reforms both reforms have already been compromised and reformers cannot back out. A realization of the state of nature of the first reform k is higher than realization $\tilde{k}$ if

$$k > \tilde{k} \rightarrow E_m O(s_{1k}, s_{2m}) \geq E_m O(\widetilde{s_{1k}}, s_{2m}) \quad (1)$$

There is a probability that the reform has lower payoff than expected. We denote this situation as $\Pr(k < \tilde{k}) \equiv \overline{Pr}$. We denote the opposite situation as $\Pr(k \geq \tilde{k}) \equiv Pr$.

Another important aspect of the model is the degree of reversibility of reforms. If the outcome of a reform is lower than the expected initially reform may not be finished. Therefore reversibility doesn´t mean to completely go back the reform but part of the investment associated to the reform can be maintained on hold. An example would be to stop the deployment of the network of an entrant service allocation when the most profitable regions are covered leaving less profitable areas uncovered because the reform is more expensive than expected due to the emergence of interference problems. We denote the degree of reversibility of a big bang reform $\xi < 0$ .In the case of the gradual approach the reversibility of the first and second reforms are defined respectively $\xi_1 < 0$ and $\xi_2 < 0$.

$$\xi \leq \xi_1 + \xi_2 \leq 0 \tag{2}$$

The payoff of a big bang reform can be written as follows

$$BB = \Pr R_2 + \overline{Pr}\ \overline{R_2} \tag{3}$$

Where

$$R_2(s_{1k}|k \geq \tilde{k}) \equiv R_2 = (1-\delta)O_m + \sigma\delta O_m \tag{4}$$

$$R_2(s_{1k}|k < \tilde{k}) \equiv \overline{R_2} = (1-\delta)O_m + \xi \tag{5}$$

$R_2$is the payoff of the big bang reform conditional to a payoff of the first reform higher than expected, $\overline{R_2}$ is the payoff of the big bang reform conditional to a payoff of the first reform lower than expected.

The payoff of a gradual reform is

$$GR = (1-\delta)H_k + \sigma\delta\overline{Pr}\,\xi_1 + \sigma\delta \Pr R_2 \tag{6}$$

Proposition 1: The condition for gradualism to dominate big band strategies under technological advance is

$$\sigma > \frac{\Pr(1-\delta)[O_m - H_k]}{\overline{Pr}\,\delta[\xi_1 - \xi]}$$

Proof see appendix

Let's consider now two different discount factors for reform 1 and reform 2 denoted by δ_1, δ_2 and an expected technological advancement after reform 2. We denote δ_0 the additional discount factor to be added to $\delta_2$if it is decided to undertake a gradual type of reform with the goal of waiting until the new technology is available. The relationship between the discount factors is:

$$1 > \delta_1 > \delta_2 > \delta_2' > 0 \tag{7}$$

$$\delta_2' = \delta_2 - \delta_0 \tag{8}$$

Under such conditions the big bang reform expression can be expressed:

$$BB = Pr[\delta_1 O_m + \delta_2 O_m] + \overline{Pr}\,[\delta_1 O_m + \delta_2 \xi] \tag{9}$$

The expression for the gradual reform is

$$GR = \delta_1 H_k + \delta_2\,\overline{Pr}\,\xi_1 + Pr[\delta_1 O_m + \sigma\delta_2' O_m] \tag{10}$$

Proposition 2: If there is a technological advancement that may enable the achievement of higher expected outcome, a gradual reform waiting to incorporate the new technology is a better strategy than a big bang reform if the additional discount factor introduced by the awaiting time fulfills the following expression

$$\delta_0 > \frac{O_m[Pr(\sigma - \delta_2) - 1] + \delta_1 H_k + \delta_2 Pr[\xi_1 - \xi]}{\sigma Pr O_m}$$

Provided that

$$\frac{O_m Pr(\sigma - \delta_2) + \delta_1 H_k + \delta_2 Pr[\xi_1 - \xi]}{O_m} > 1$$

Proof see appendix

5. Conclusions

Spectrum management reform can be analyzed from the perspective of the political economy of reform adding the idiosyncrasies of high technology industries. Such industries are characterized among other factors by intense technological eveolution. Another distinctive feature is the impossibility of completely reverse reforms. Reversibility here does not mean to return to the status quo but to leave unfinished some reform components. For example, the result of an unfinished reform may be network deployment solely in profitable areas, enabling the benefits only in high densely populated areas, or in other cases, the expected number of transactions may not be achieved, being reduced to the cases where uncertainty is the lowest. An Additional characteristic is that the reformers of spectrum management cannot usually choose the order of reforms. The reason is that reforms must not only be complementary but also have to be carried out in a particular order. An example of a forced sequence of reforms is the adoption of a new technology, followed by a new spectrum allocation and a subsequent reallocation of the newly freed up spectrum to the incumbent and to an entrant user. This reform package must follow a particular order. The reformer instead must choose the right moment to start the reform and the implementation strategy, either big bang or gradual.

We offer an expression to determine when to choose a gradual or big bang reform depending on current technology and an expression to determine whether to wait or not for a new technological advancement. The results depend on the expected outcome of the reform, the probability of realization of the states of nature and the cost of reform reversal.

We have also shed some light on the types of spectrum management reform by categorizing them in terms of the type of efficiency gain that the reform is intended to produce,- productive, allocative and dynamic-. Furthermore, we have analyzed market failures responsible for uncertain outcomes to help tackling the underlying causes of failure.

The use of market mechanisms to improve productive efficiency, -intra-band technological change, auctions and secondary market transactions-, has already been fruitful on the portions of spectrum traditionally been used for commercial purposes. In such cases, market forces may result in productively efficient transactions provided that Coasian schemes are fulfilled; low transaction costs and well-defined property rights.

Transactions enabling allocative and dynamic efficiency are not yet possible to be produced solely by market forces. New approaches such as the design of incentive auctions give a broader role to market participants but the active contribution of the regulatory authority to design the direct and reverse auction and the repackaging process is still necessary. A different instrument, the authorization of new technologies enabling to share spectrum previously used by an incumbent service, is improving allocative efficiency making market agents more involved in spectrum allocation. However, sharing techniques are usually limited either to short range devices, the case of underlay spectrum techniques, or by the availability of unused spectrum in certain areas or time slots, the case of dynamic spectrum sharing. Spectrum sharing techniques that are useful to improve the efficiency of spectrum allocation cannot be used as means to enable the change of use of a band solely by market forces.

An additional innovative approach to spectrum management is the definition of spectrum licenses in terms of the maximum acceptable interference that devices should be able to accept. This technique may pose a significant advancement to achieve the goal of enabling market transactions changing the allocation of spectrum. However, better knowledge of interference problems and more importantly the design of new types of licenses, would be required to move forward.

The efficiency of changes of spectrum allocation by market forces may be hampered by the existence of externalities associated to services and services that are pure public goods. In addition, market agents may not be willing to change to a more socially efficient spectrum allocation due to several reasons including the existence of uncertain costs but also because current market structure may grant market power to incumbent users, a situation that they are interested to preserve. Regulators must analyze the possibility to enable the re-allocation of spectrum by market forces on a case by case basis.

Annex 1: Proof of propositions

Proposition 1

According to (3) and (6), for a gradual to be superior to a big bang reform, the following must be fulfilled

$$(1-\delta)H_k + \sigma\delta\overline{Pr}\,\xi_1 + \sigma\delta \mathrm{Pr} R_2 > Pr[(1-\delta)O_m + \sigma\delta O_m] + \overline{Pr}\,[(1-\delta)O_m + \xi]$$

$$(1-\delta)H_k + \sigma\delta\overline{Pr}\,\xi_1 > Pr[(1-\delta)O_m + \sigma\delta O_m] + \overline{Pr}\ \sigma\delta\xi$$

$$\sigma\delta\overline{\mathrm{Pr}}[\xi - \xi_1] \ > Pr[(1-\delta)O_m - (1-\delta)H_k]$$

$$\sigma > \frac{\mathrm{Pr}(1-\delta)[O_m - H_k]}{\overline{Pr}\,\delta[\xi_1 - \xi]}$$

We know that $O_m > H_k$ and $\xi \leq \xi_1 \leq 0$ and δ <1 therefore all the terms of the expression are positive

Proposition2

According to (9) and (10), for a gradual to be superior to a big bang reform, the following must be fulfilled

$$\delta_1 H_k + \delta_2\,\overline{Pr}\,\xi_1 + Pr[\delta_1 O_m + \sigma\delta_2' O_m] > Pr[\delta_1 O_m + \delta_2 O_m] + \overline{Pr}\,[\delta_1 O_m + \delta_2\xi]$$

$$\sigma Pr\delta_2' O_m > \ \overline{Pr}\,\delta_1 O_m - \delta_1 H_k + \delta_2 Pr(\xi - \xi_1) + Pr\delta_2 O_m$$

$$PrO_m[\delta_2(\sigma-1) - \sigma\delta_0] > \ \overline{Pr}\,\delta_1 O_m - \delta_1 H_k + \delta_2 Pr(\xi - \xi_1) + Pr\delta_2 O_m$$

$$\delta_0 > \frac{O_m[Pr(\sigma-\delta_2) - 1] + \delta_1 H_k + \delta_2 Pr[\xi_1 - \xi]}{\sigma Pr O_m}$$

For this expression to be positive it must be fulfilled the following condition

$$O_m[Pr(\sigma-\delta_2) - 1] + \delta_1 H_k + \delta_2 Pr[\xi_1 - \xi] > 0$$

$$\delta_1 H_k + \delta_2 Pr[\xi_1 - \xi] > O_m[1 - Pr(\sigma - \delta_2)]$$

$$\frac{O_m Pr(\sigma-\delta_2) + \delta_1 H_k + \delta_2 Pr[\xi_1 - \xi]}{O_m} > 1$$

References

Alden, J. (2012). Exploring the value and economic valuation of spectrum. Report prepared for the ITU.

Baumol, W., and Robyn, D. (2006). Toward an Evolutionary Regime for Spectrum Governance: Licensing or Unrestricted Entry? Brookings Institution Press.

Bazelon,C., McHenry,G.(2013). Spectrum value. *Telecommunications Policy*, Volume 37, Issue 9, October 2013, Pages 737-747, ISSN 0308-5961.

Benkler, Y. (2002). Some Economics of Wireless Communications. *Harvard Journal of Law & Technology*, 16 (1), 25-83.

Cave, M (2002). Review of radio spectrum management. Report for Department of Trade and Industry and HM Treasury.

Cave, M. and Webb, W. (2003). Designing Property Rights for the Operation of Spectrum Markets. University of Warwick (UK), Warwick Business School Papers in Spectrum Trading, num. 2.

Cave, M., Doyle, C. and Webb, W. (2007) Essentials of Modern Spectrum Management. Cambridge University Press. ISBN 9780511536724

Cave, M., Webb,W. (2012).The unfinished history of usage rights for spectrum. *Telecommunications Policy*, 36(4), 293–300.

Cave,M. Pratt,N. (2016). Taking account of service externalities when spectrum is allocated and assigned, *Telecommunications Policy*, Volume 40, Issues 10–11, October 2016, Pages 971–981

Coase, R.H. (1959). The Federal Communication Commission. *The Journal of Law &Economics*, 2, 1-40.

Coase, R.H. (1960). The Problem of Social Cost. *The Journal of Law & Economics*, 3 (1), 1-44.

Cornes, R. & Sandler, T. (1996). The Theory of Externalities, Public Goods and Club Goods. Cambridge University Press.

Dewatripont, M.,and Roland,G.(1995).The design of reform packages under uncertainty. *The American Economic Review*, 85(5), 1207–1223.

Dixit,A. (1992). Irreversible investment with uncertainty and scale economies. LSE Research Online Documents on Economics 19372, London School of Economics and Political Science, LSE Library.

Dixit, A. and Pindyck, R. (1994). Investment under uncertainty. Princeton, NJ: Princeton University Press, 1994.

European Commission (2012). Communication from the Commission to the European Parliament, the Council, the European economic and social Committee of the regions. Promoting the shared use of radio spectrum resources in the internal market.

FCC (2012). Notice of Proposed Rulemaking and Order in the matter of amendment of the Commission's rules with regard to commercial operations in the 3550-3650 MHz band, 12 December 2012.

Frias, Z., González-Valderrama, C., Pérez-Martínez, J. (2016). Assessment of spectrum value: The case of a second digital dividend in Europe. Telecommunications Policy. ISSN: 308-5961; http://dx.doi.org/10.1016/j.telpol.2016.12.008

García, J.M. Valiño, A. (2013). Propuestas para la evaluación de la eficiencia de una atribución de espectro a un servicio de radio-comunicaciones. GCG: Revista de Globalización competitividad y Gobernabilidad. Vol7. Num2, 85-105.

García-Zaballos, A., Foditsch, N. (2015). Spectrum Management, the key lever for achieving universality. Inter-American Development Bank. IDB-MG-322.

Gillet, S. E., Lehr, W.H., Osorio, C.A., Sirbu, M.A. (2006). Measuring broadband's economic impact. Final Report prepared for the U.S. Department of Commerce, Economic Development Administration.

Goss, E. (2001). The Internet´s contribution to U.S. productivity growth. *Business Economics*, 36 (4), 32-42.

Hazlett, T.W. (1998). Spectrum Flash Dance: Eli Noam's Proposal for 'Open Access' to Radio Waves. *Journal of Law & Economics*, 41 (2) 805-820.

Hazlett, T.W. (2003). Liberalizing US Spectrum Allocation. *Telecommunications Policy*, 27 (7) 485-499.

Kwerel, E. & Williams, J. (2002). A Proposal for a Rapid Transition to Market Allocation of Spectrum, Washington, D.C., FCC - Office of Plans and Policy, Working Paper, 38.

Kwerel, E., Williams, J. (2010).The spectrum regulator's dilemma in a dynamic world: Limiting interference without stifling innovation. Conference held at the Information Technology and Innovation Foundation (ITIF), Washington, DC.

Levin, H. J. (1970). Spectrum Allocation without Market. *The American Economic Review*, 60(2), 209-18

Litan, R. & Rivlin, A. (2001). Projecting the economic impact of the Internet. *The American Economic Review*, 91 (2), 313-31.

López Sánchez, (2004). ¿Pueden las Tecnologías de la Información mejorar la Productividad? *Universia Business Review*, 1(1), 82-95.

López Sánchez, J.I., Minguela Rata, B., Rodriguez Duarte, A., Sandulli, F.D. (2006). Uso de Internet y paradoja de la productividad: el caso de las empresas españolas. *Cuadernos de Economía y Dirección de la Empresa*, 2, 149-174.

Minervini, L.F. and Piacentino (2007). Spectrum management and regulation: Towards a Full Fledged market of spectrum bands. Working Paper num. 7. Università degli Studi di Macerata. Dipartamento di Studi Sullo Sviluppo Economico.

Minervini,L.F. (2014) Spectrum management reform: Rethinking practices. *Telecommunications Policy*, Volume 38, Issue 2, March 2014, Pages 136-146, ISSN 0308-5961.

Noam, E. (1998) Spectrum Auctions: Yesterday's Heresy, Today's Orthodoxy, Tomorrow's Anachronism. Taking the next step to open spectrum access. *Journal of Law and Economics*, 44, 765-790.

PCAST (2012). United States President's Council of Economic Advisors on Science and Technology (2012). Report to the President. Realizing the Full Potential of Government-Held Spectrum to Spur Economic Growth, July 2012.

Pindyck, R.S. (1993). Investments of uncertain cost. *Journal of Financial Economics*, num. 34, Issue 1, 53-76.

Prasad, R. (2015). The production function methodology for estimating the value of spectrum, Telecommunications Policy, Volume 39, Issue 1, February 2015, Pages 77-88, ISSN 0308-5961

Radio Spectrum Policy Group (2011). RSPG Report on Collective Use of Spectrum and other spectrum sharing approaches.

Smith-Nera (1996) Study into the use of spectrum pricing. London: Radio communications Agency, UK.

International Telecommunication Union (2009). ICT regulation Toolkit. Module 5. Radio Spectrum Management.

Webb, W (2009). An optimal way to licence the radio spectrum. *Regulatory and Economic Policy in Telecommunications*, 2, mar09, 100-112.

Werbach, K. (2004). Supercommons: Toward a United Theory of Wireless Communication. Texas Law Review, 82 (4), 863-973.

The World Bank Group (2016). World Bank Group. 2016. World Development Report 2016: Digital Dividends. Washington, DC: World Bank. License: CC BY 3.0 IGO

Notes

[1] AIP have been implemented in the UK and Canada. https://www.ofcom.org.uk/consultations-and-statements/category-1/aip/update201008

[2] More information about the process of an incentive auction can be found at: https://www.fcc.gov/wireless/auction-1000

[3] Normally there will be a different limit for each type of device, for example, one for transmitters and a different one for receivers.

[4] In some cases the establishment of ex-ante regulation such as the definition of spectrum caps might be required to achieve effective competition. Spectrum caps are limits to the maximum amount of spectrum that a firm is allowed to use to provide a service. In other occasions ex-post measures might be enough.

[5] A complementary explanation is that digital radio receivers are still expensive, and the use of analogue devices is highly extended.

[6] Examples of infrastructure involved in high sunk costs are the deployment of satellite networks or undersea cables.

[7] Profit generated on account of the allocated spectrum.

[8]A list of Federal Spectrum uses elaborated by NTIA (National Telecommunications and Information Administration) can be found at https://www.ntia.doc.gov/files/ntia/publications/spectrum_use_summary_master-07142014.pdf

[9] Roaming is the possibility to continue receiving a service when travelling abroad. Roaming can be enjoyed either because a frequency band is harmonized or because user´s terminal can operate in the different frequency bands used in each country.

[10] This includes reusing infrastructure like telecommunication towers, the electric feeding system, communication devices and also rights of way and site leasing.

[11] There are three types of interference: geographical, out of band and in the band itself. Licenses would establish the maximum interference from the adjacent channel and the maximum out of band interference (Cave and Webb, 2003).

[12] The same radio communication device can work as a transmitter and receiver. For example, a mobile terminal is able to transmit and receive.

[13] Unwanted energy can also be the result of a combination of emissions, e.g. intermodulation products, or inductions upon reception.

[14] For example the use of transmission power control techniques intended to increase mobile phones battery life increases the probability of harmful interference to other systems http://www.erodocdb.dk/docs/doc98/official/pdf/ECCRep138.pdf

[15] Transmission devices send part of the power emitted outside their transmission band. The amount of power sent out of band depends on the features of the transmitter´s output filter. Furthermore, due to the existence of intermodulation products, interferences can be produced on frequencies different to those emitted, due to the emission of a mixture of frequencies taking place in the transmitter. There also exists interference caused by the accumulation of out-of-band signals from multiple nearby transmitters.

[16] When the transmitter of a new service is located at the coverage edge of the existing services the probability of interference increases. Services with similar typical power tend to share the same or nearby sites thus, reducing the possibility of coverage holes.

[17] A receiver can be blocked in the presence of a nearby high power signal. Under such condition the receiver is unable to receive any signal.

[18] The existence of guard bands may be necessary even to expand the allocation of a service to adjacent bands. For example, the expansion of mobile broadband to adjacent frequency bands requires creating a guard band of unused spectrum.

[19] For example, freeing up the so called digital dividend frequencies required to simulcast new and old television channels during a period of time to let people adapt building reception facilities.

CHAPTER 3: PRICE DISCRIMINATION OF OTT PROVIDERS UNDER DUOPOLISTIC COMPETITION AND MULTI-DIMESIONAL PRODUCT DIFFERENTIATION IN RETAIL BROADBAND ACCESS

Abstract: Network neutrality regulation prevents access providers to price discriminate between content providers and product differentiation in terms of connection quality in the retail broadband access market. This paper analyzes the economic implications of price discrimination under duopolistic competition and multi-dimensional product differentiation in retail internet access using a sequential-moves game theoretic model. Under this framework, we discuss the impact of product differentiation and price discrimination on social welfare, and offer systematic simulations using feasible ranges for parameters value to help discern the impact of departing from network neutrality regulation on social welfare.

Keywords: network neutrality, two sided markets, price discrimination, product differentiation, competition policy.

JEL Classification Codes: C70, D43, L51, L52, L86.

1. Introduction

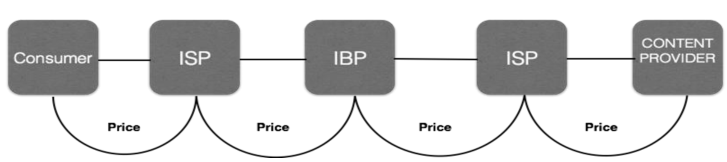
The internet consists of a set of networks intended to the transportation of data packages. These packages are the unit item to provide a myriad of different services and applications such as e-mail, web browsing, voice communications, search services and linear and on-demand video distribution. The inventory of services delivered over the top (OTT) of internet networks is not a closed list, but rather a fast evolving catalogue with increasingly innovative applications and services appearing constantly.

Since the conception of the internet regulation has preserved the principle that all data packages must be treated equally regardless the type of digital service being provided to internet users. This imply that networks are agnostic with respect to the OTT services been offered. Thus, there is not network support or aid to OTT service provision in terms of guarantying speed, delay or delay variability of packages. Furthermore, the slow down or blocking of a particular kind of internet traffic is forbidden. This regulatory principle is known as the network neutrality regulatory approach to internet access.

From the economic standpoint, network neutrality means that internet providers are only enabled to charge consumers and content providers for connectivity and are not allowed to exert price discrimination to content providers regardless the profits they enjoy or the congestion they generate in the internet consumer’s local access network. It is neither possible for connectivity providers to differentiate the internet access product in terms of quality of connection, defined as higher quality the lower the delay.

Network neutrality is a regulatory approach to internet access pricing by which business and residential internet consumers and content and application providers (CP) pay the price to access the internet access to their respective local Access Provider[1] (AP). Additionally, there are payments among other internet network segment owners that are necessary to transport CPs´ traffic to internet users. APs only own or use one of the internet network segments, the access network, but that it is not enough to reach all the content and applications available on the Web. Consequently, APs need to interconnect to other network suppliers, e.g. backbone network providers, to be able to offer end to end communications. There are economic transactions among APs and Backbone Providers (BP). Backbone Providers either exchange traffic among them and pay an interconnection price to other BPs or agree to share their networks if net payments between them approximately compensate. Under network neutrality all payments are made linearly, therefore, the internet consumer’s local AP does not receive any payment from CPs generating traffic in a remote AP network; see Figure 3.1.

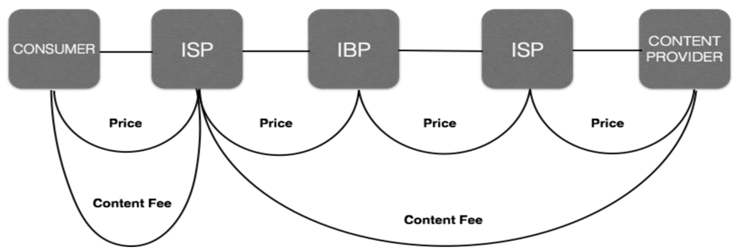
Figure 3.8: Network Neutrality Approach



Authors based on Economides and Tag (2012)

If the consumer's local AP is enabled to charge a fee either to consumers or more probably to CPs to differentiate connection quality, thus discriminating their ability to pay, the network neutrality paradigm would be broken. Under this approach, that Wu and Yoo (2006) denominated "network diversity", traffic is not treated equally anymore and consumers and/or content and application providers could be charged for different qualities of internet connectivity; see figure 3.2.

Figure 3.9: Network Diversity Approach



Authors based on Economides and Tag (2012)

There are multiple interrelated players and industries that are part of the digital goods and services provision value chain. Examples of such industries are content and application providers, device manufacturers, information systems designers, mobile and fixed internet network providers and digital advertising companies. Players include governments, regulators, the academia, ICT companies and citizens. A framework describing the internet ecosystem can be found in García-Zaballos and González-Herranz (2013).

Content and application providers argue that network neutrality is the cornerstone for an open internet and has been responsible for the successful emergence of innovative applications and OTT services such as voice or video over IP. However, APs respond that the investments required to expand the access network and to adopt next generation technologies are threatened by network neutrality regulations. Opponents of network neutrality also maintain that the development of new technologies such as some applications of the internet of things would be endangered by network neutrality regulation.

From the APs business model perspective, two forces are responsible for their request to abolish the network neutrality principle. Traffic demand is increasing exponentially, congesting fixed and especially mobile networks. This surge in demand is the result of the emergence of new devices such as tablets and smartphones, and the increasing usage of multimedia applications. For example, Netflix and Google are responsible for about 50% of downstream traffic in fixed networks during peak periods in the US[2]. There is also a shift underway in the balance among the agents in the value chain, away from the predominance of connectivity providers towards the increased importance of CPs[3]. In this context, the economic analysis of the different regulatory models for internet access is of utmost relevance for the expansion of internet networks and services.

APs claim that the value proposition underlying the creation of prioritized lanes for internet traffic is the reduction of delay and congestion, but with current technology, prioritization might be trying to solve technical problems that currently don´t exist. New technological developments such as protocols that reduce congestion, internet exchange points and peering techniques reduce data packages delivery cost, packages losses and latency and improve routing efficiency. After years of innovation in the delay effort-less internet, it may be too late for APs to argue for a dedicated path as a solution for congestion and delay. Only very critical jitter and latency applications such as telemedicine may need a dedicated path and reserved internet node capacity.

2. Related literature

The impact of internet access regulation has been analyzed from different points of view. It has being scrutinized from a view of achieving different public policy objectives such as income redistribution and inequality reduction, ensuring freedom of speech and expression and guaranteeing consumer’s data privacy. A different perspective is to examine the topic from the competition policy angle by evaluating the impact of changes in regulation to effective competition, e.g. the effect of vertical and horizontal integrations in the value chain of internet service provision encouraged by regulatory modification. Finally, it has been studied from an economic efficiency perspective by calculating social welfare and by estimating the incentives to invest and innovate in internet network deployment and in the creation of new OTT services. The most efficient social solution is analyzed from the productive, Pareto and dynamic standpoints. The following paragraphs analyze existing material on these three ways of approaching the network neutrality debate, though this paper focus the analysis on the impact of internet access regulation on social welfare.

One public policy objective is income redistribution. Examples of this kind of policy in the ICT industry are the obligations that telephone and internet access providers are required to comply with, like emergency call services, the rural and remote areas coverage requirements or the contributions to universal service funds. The asymmetric regulation of OTT and internet access services may pose concerns related to the different redistribution policy burden borne by companies providing similar services[4]. Based on this rationale, internet access providers claim for a level playing field. Notwithstanding, we don’t center the analysis on this topic since symmetric regulation may co-exist with network neutrality regulation.

Another public policy objective is the safeguard of citizen’s freedom of expression and data privacy. An open internet has been regarded as a pre-requisite to safeguard free speech and privacy. Without network neutrality, APs may inspect consumer’s information or block some contents, thus reducing freedom of speech[5]. However, regulations limiting such practices can co-exist with the existence of priority internet traffic classes.

In reference to internet regulation impact on competition, existing literature studies agents’ incentives to undertake anticompetitive behavior. Hemphill (2008) enumerates three network diversity potential constraints to effective competition and analyses whether network neutrality is a correct regulatory response. The first constraint is the potential market exclusion of CPs through abuse of dominance after the vertical integration of an AP with significant market power and a rival CP. Hemphill maintains that network neutrality is an unnecessary rule since antitrust laws already deal with the problem, e.g. the "Microsoft-Netscape[6]"case. However, this case might also exemplify how lengthy court proceedings may allow a monopolist to keep its market power endlessly.

The second is the potential market exclusion of a CP through refusal to deal. An AP might have an incentive to block CP's services that substitute legacy telecommunication services such as voice, text or videos, e.g. the case "AT&T versus Vonage[7]" in the US. Hemphill holds that a sound antitrust regulation would solve this problem with the only exception of exclusion of social production, e.g. Wikipedia. A fee for prioritized traffic may drive Wikipedia out of the market even if social production of online encyclopedias is more efficient than market production.

Finally, a third problem, is the reduction of CP's incentives to invest and innovate if APs charge a fee to CPs. Low investment incentives could also lead to reduced competition in the long run. Hemphill points out that internet access regulation is a joint problem since not only innovation in OTT services should be taken into account but also innovation in broadband infrastructure deployment.

The innovation dilemma is also analyzed by Van Schewick (2006) and Wu and Yoo (2006). Van Schewick states that keeping innovation in content creation through network neutrality brings wider benefits than costs. Wu explains that the key element to understand the impact of regulatory change is the economics of deployment of the last mile internet access network infrastructure. Abandoning network neutrality regulation would harm innovation and investment in OTT services provision in exchange for very limited help to solve the last mile infrastructure problem because of the existing lack of effective competition among internet access providers. However, Yoo considers that the spectrum-based internet access technologies are changing the economics of the last mile network, making competition feasible and the internet access market contestable. Under such assumption, changing network neutrality regulation would be an effective measure to solve the last mile network deployment problem. Earlier, Huergo (2004) had also identified the positive effect on competition of new technologies increasing efficiency of spectrum use. Wright (2013) and Faulhaber (2011) point out that the elimination of network neutrality regulation should be subject to the rule of reason. Under this rule, the existence of prioritization should not be categorically prohibited in the absence of a demonstrated harm to competition. For Shin (2014), an effective regulatory strategy should include ex-ante principles, guidelines for acceptable practices and precompetitive regulations.

An additional AP's practice that should be analyzed from the competition perspective is the so called zero rating. It consists on the exemption of certain favored internet content from internet access users' monthly data caps. The potential adverse effect to effective competition of such practice is similar to the exclusion of a CP happening when there is a vertical integration of an AP and rival CP. There is increasing support to ban or at least regulate this practice[8]. In developing countries, especially those where users have very low ability to pay, this practice put the regulator in the dilemma of choosing between potential competition distortion and unaffordable access to internet content for most of the population.

Theory of strategy may help us to obtain some qualitative intuition to understand market agents' behavior towards competition. In this context, it is useful to use Porter's (1991) five forces strategic framework to understand the impact of departing from network neutrality on the different industries involved. The application and CP's industry would face decreased threat of new entrants and a lower threat of substitute products since a fee to content provision constitutes a new entry barrier. This industry would face a similar rivalry among competitors and bargaining power of buyers and highly increased bargaining power of suppliers. As a result CP's industry would probably be less attractive with network diversity.

The AP's industry would face reduced bargaining power of some buyers -the CPs-, a very limited increased threat of new entrants due the existence of high sunk costs in network deployment, a fall in the threat of substitute products or services and increased own bargaining power. Bargaining power of suppliers and rivalry among competitors would remain the same as with network neutrality. As a result AP's industry would probably be more attractive with network diversity.

As we have seen, there exists diverse regulatory responses limiting anticompetitive behavior, for this reason, we consider that other regulatory gaps being filled, internet access regulation turns to be an economic efficiency maximization problem. Extant economic modeling approaches internet access regulation carrying out two non-mutually exclusive types of analysis. One approach is to examine how a regulatory change would affect social welfare in the short run, a different one is to analyze the effects of regulation on the investment incentives of CPs and APs in the long run. Network neutrality impact on social welfare is analyzed using two main types of methodologies. One approach is to use two-sided market models looking at cross externalities between the two sides of the market and equilibrium prices. Related literature can be found in Economides and Tag (2012), Armstrong (2006), Rochet and Tirole (2003, 2006), Nocke et al. (2007), Weyl (2010), Mialon and Banerjee (2013), Jay, Jeon and Kim (2014), Bourreau et al (2015). A second approach is to focus on measuring the impact of congestion management in total welfare in a context of bandwidth scarcity. This perspective is offered in Economides and Hermalin (2012), Hermalin and Katz (2007), and Cheng et al. (2006).

Two-sided market research shows different conclusions. Armstrong (2006) found that the key determinants of equilibrium prices in a two-sided platform are the magnitude of the cross-group externalities, whether fees charged by the platform are lump-sum or per-transaction, and whether the agents have multi-homing possibilities[9]. Rochet and Tirole (2003, 2006) and Mialon and Banerjee (2013) maintain that the structure of the two-sided market is an important element to determine total welfare. The latter show that vertical relations among content and internet APs are a key element to calculate total welfare. On this same line of thought Jay, Jeon and Kim (2013) note that the impact of network neutrality crucially depends on the CP's business model. Weyl (2010) offers a model in which users are heterogeneous in two dimensions, income and scale. Economides and Tag (2012) analyze the impact of regulation on network effects to determine the equilibrium prices charged by the platform and the resulting number of consumers and CPs. They find that for most parameter values, total surplus is higher at zero fee, thus, with network neutrality.

Regarding congestion models Hermalin and Katz (2007) find that restricting a monopolist AP to have a single not differentiated product has the following effects: application providers looking for the low quality variant are excluded from the market, intermediate users enjoy higher and more efficient qualities than demanded, and top users suffer lower and less efficient qualities than desired. Total surplus may rise or fall depending on the parameters used. The extension to the duopolistic model show similar welfare effects. Cheng et al. (2007) use a game theoretic model approach in which congestion can be avoided paying for preferential access. They find that departing from network neutrality would benefit APs and hurt CP's profit. Economides and Hermalin (2012) assume network congestion and find that network neutrality increases total welfare if the elasticity of content demand with respect to transmission time does not increase with households' time sensitivity for content.

Several studies analyze the impact of internet access regulation on the investment incentives of CPs and APs. Choi and Kim (2010), Cheng, Bandyopadhay and Guo (2011), Economides and Hermalin (2012, 2015), and Njoroge et al (2013) analyze AP's incentive to invest in network expansion and/or network upgrade to reduce congestion. Choi and Kim (2010), Economides and Hermalin (2012), Bourreau et al (2015) study CP's incentives to enter the market or to invest in the creation of new contents and applications. Peitz and Schuett (2015), and Choi, Jeon Kim (2014) investigate CP's incentives to invest in compression technologies to reduce congestion in fixed and mobile access networks.

There also exists, although scarce, empirical literature on network neutrality. Ford et al (2010) study the impact of network neutrality in broadband deployment in rural areas and conclude that neutrality reduce network expansion to high-cost deployment rural areas. Balezon (2010) analyzes the topic from the perspective of job creation in the US and finds that neutrality might pose a negative impact on broadband sector jobs. Clarke (2009) use simple quantitative modeling to show that neutrality may hinder the satisfaction of future video and other high bandwidth applications demand. Lee et al (2011) study four types of CPs with different sensitivity to traffic delay and measure network usage efficiency using a Tobit regression model. They find that establishing traffic classes don't significantly decrease CP's efficiency.

The network neutrality dilemma has also entered the political agenda in the European Union and the US. In the US, new regulations[10] have been passed after a long and tortuous political process. In the European Union, open Internet regulation has also been reviewed. Both economic areas recognize the importance of avoiding blocking and the slowing down of any digital service to guarantee competition, freedom of expression and consumer's privacy. Both have passed regulation to impose network management transparency obligations to APs. However, there is a central point where approaches differ. US regulation prohibits the existence of paid prioritization of any kind whereas the European regulation has allowed for a different treatment specialized services provided that these services do not harm the availability and quality of the open internet access. Examples of specialized services are IPTV, high definition videoconferencing, automated driving or healthcare services like remote surgery. National regulatory authorities have the obligation to examine traffic management practices to ensure that minimum quality of service requirements are fulfilled and open access is not damaged. In case of infraction, Member States would set effective, proportionate and dissuasive penalties.

The aim of this paper is to study the economic efficiency of internet access regulation, assuming other regulatory concerns solved. In this sense, it is important to point out the difference between completely unregulated control of data packages and the allowance of prioritized traffic under regulatory control and open internet quality of service requirements. The first may pose risks to competition, freedom of expression and privacy whereas the second is an economic efficiency issue.

3. The model

We move the network neutrality debate to the issue of product differentiation of retail internet access in a multidimensional space and price discrimination among content providers. Our setup departs from the model described in Liu and Shuai (2015) -multidimensional product differentiation and endogenous and fixed firms' location- to subsequently determine the equilibrium price of internet access and the quantity of internet content generated by CPs under different scenarios. We consider three potential scenarios and calculate total welfare for each; network neutrality (NN), network diversity (ND) operated by all APs, and ND used by one AP whereas the other continue acting under NN. The selection of one of the two ND scenarios might happen either as a result of the strategic behavior of APs or as a technology or market constraint in areas with very low population density[11]. The elaboration of digital content is financed through internet advertising but we can expect similar results for content financed by internet consumers. The model is a sequential-moves game. APs move to a limited new well known locations when there is a change in regulation from NN to ND. The sequential-moves game stages are the following:

1. In the first stage APs select the internet access price. We calculate the equilibrium in all three possible scenarios assuming duopolistic Bertrand's competition and the existence of a unique pure strategy Bertrand-Nash equilibrium. We also analyze in this stage the impact of a change in unit transport costs on the equilibrium price of internet access and on AP's market share. The equilibrium is defined by the quadruplet $({p_A}^*, {p_B}^*, {q_A}^*, {q_B}^*)$. A summary of all the notations used in this paper is provided in Table 3.1.

2. In the second stage, CPs choose the maximizing quantity of content to be provided $c^*({q_A}^*, {q_B}^*)$ in order to maximize their profit given the APs market shares calculated in the first stage. In the scenarios for which there exists prioritization, APs also choose simultaneously in this stage the optimal fee per unit of content per user to be charged to CPs.

We make five important market structure assumptions. First, firms have fixed location in their differentiation strategy because of existing regulation, technology constraints and market features such as the existence of economies of scope and other market structure limitations[12]. Whereas most of product differentiation literature focus on how firms locate using a first location then price type of game, Hotelling (1929), D'Aspremont et al (1979), Economides (1986) etc… We focus the analysis on the impact of unit transport costs on the equilibrium price of internet access. Second, in cases where the AP is allowed to offer prioritized traffic and prioritization is technically feasible, we assume that CPs will be price discriminated in such a way that all CPs whose service quality depend on delay or delay variability will pay for prioritization; the remaining CPs will provide their services using non-prioritized traffic regardless the existence or not of priority traffic for other services. We make this assumption based on the following rationale. Without the possibility of prioritization, CPs offering delay sensitive services have been responsible for reducing delay and delay variability using techniques such as internet exchange points, peering or protocols that reduce congestion. However, the existence of traffic prioritization may pose an additional obstacle to the provision of a service with sufficient quality because of the increased network congestion if the CP does not pay for the prioritized traffic, thus the CP will pay for prioritization. Third, APs will charge the cost of prioritization only to CPs, thus, internet consumers will not be charged. We support this assumption in the two sided market features, for example because of the limits to impose payments only to consumers that use services that are susceptible of being prioritized, different demand elasticity in both sides of the market and the possibility to attract more internet consumers if they are not charged for prioritization, which benefit both APs and CPs. For similar reasons, we also consider that CPs are neither able to nor willing to pass this costs to internet consumers. These assumptions are common in the literature; see Rochet and Tirole (2006). The existence of prioritized traffic is regarded by consumers as an internet access product differentiation feature. Fourth, we assume internet access duopolistic provision; such assumption is underpinned by the existence of high sunk costs in network deployment that hinders market entry. Fifth, we assume for all scenarios that there is regulatory control over service blocking in order to prevent competition distortion and forced demand for priority that otherwise would not be requested. Sixth, CPs are independent monopolist in its own market, thus assuming a winner takes all type of content provision market. This assumption is also common in the literature; see for example Economides and Tag (2012). The analysis with different imperfect competition content provision market structures is left for future research. It is also important to note that the potential impact of departing from network neutrality regulation on investment incentives is also left for future research. We analyze the advisability of departing from network neutrality regulation in a context of duopoly in the market of internet access and horizontal product differentiation in a two dimensional space. One product differentiation feature is related to the preference that consumers place to delay or the delay variability of data packages, therefore linked to network neutrality regulation. Some consumers are reluctant to traffic prioritization because they use internet services for which prioritization cannot be appreciated. A network with prioritized traffic may increase congestion and reduce network capacity available for services that don't require prioritization. Other group of users would be indifferent to prioritization because they neither use services susceptible to be prioritized nor require high internet capacity. Finally, a third group would prefer prioritization because they use a set of internet services that include those sensitive to traffic delay or delay variability. Furthermore, APs will not be willing to charge internet consumers a premium price for prioritization because of the reasons explained in the third assumption of the previous paragraph. The other product differentiation feature models consumer preferences over a set of options offered by the AP to internet users, such as the set of other products bundled to internet access (e.g. fixed telephony, television, mobile data and voice…). Some consumers prefer to pay for a single product, broadband, so that they are not forced to pay for the other elements of the bundle that they don´t need, whereas others are willing to pay for the whole package. The latter are interested in all the products because the price of the bundle is usually lower than the sum of the individual prices due to the existence of economies of scope in the production of the components of the package. A more detailed explanation of the differentiation dimensions can be found in section 3.1. The consideration of two-dimensional product differentiation is also a contribution to existing literature on network neutrality since existing studies only contemplate a single dimension.

Table 3.5: List of notations[13]

Notation	Description
U_A , U_B	Utility of a consumer that buy internet access either from AP A or B
v	Intrinsic value of internet connectivity for consumers
α	Ratio of content to the maximum content under perfect competition in each scenario
β	Proportion of prioritized content to total content in the market equilibrium
$w, \overline{w}$	Waiting time with and without priority traffic
d	Marginal value that a consumer places on internet access bundled products per unit
t	Marginal value that a consumer places on internet access waiting time per unit
x	Dimension of product differentiation in terms of bundled products options
y	Dimension of product differentiation in terms of waiting time
c, c_p, c_{np}	Quantity of content of a CP, content susceptible of prioritization and non-prioritized content
C, C, C_{np}	Total quantity of content generated, and total prioritized and non-prioritized content
r_i, r_i^c, r_i^a	Relevance of a particular content, and relevance for internet consumers and advertisers
x_1, x_2	Marginal consumers indifferent to connect to either AP, touching the square horizontal lines
a_i	Price per user of the minimum amount of content of an internet service
b	Slope of the internet advertisement inverse demand curve
N^{cp}	Total number of content providers
n_p^{cp}, n_{np}^{cp}	Number of providers that would and would not pay for prioritized traffic
p_A , p_B	Price of internet access of AP A and B
$p^a(c)$	Price of advertising per unit of content per internet user
$g(c)$	Function describing the fee charged to a CP in network with prioritization
T	Marginal fee charged to a CP per unit of content
$L_{1A}, L_{2A,} L_{1B}, L_{2B}$	Position of AP A and AP B in the product differentiation square
$F_1\ (x), F_2\ (y)$	Distribution functions of consumers in dimensions x and y
$f_1(x)$, $f_2(y)$	Density functions of consumers in dimensions x and y
q_A , q_B	Market share of AP A and AP B
$\pi_A, \pi_B, \widetilde{\pi_A}, \widetilde{\pi_B}$	Profit of AP A and AP B in the neutral and discriminatory networks
PS_{AP}	Total producer surplus of APs
$\pi_{np}^{cp}, \pi_p^{cp}, \tilde{\pi}_{np}^{cp}, \tilde{\pi}_p^{cp}$	Profit of a CP in the neutral and discriminatory network
$CS, CS_{p,} CS_{np,}$	Total surplus of advertisers, and surplus with prioritized and non-prioritized traffic
$cs_{p,} cs_{np,}$	Surplus of an advertiser with prioritized and non-prioritized traffic
CS_A, CS_B	Consumer Surplus of internet consumers connected to AP A or APB
TW	Total welfare

3.1 Access Provider's location, product differentiation strategy and market share

In our model there are two product differentiation dimensions. The x dimension is related to a set of products bundled to internet access. Providers can bundle internet access with a bunch of other different services like mobile internet access, IP television, mobile voice communications, fixed voice communications and video on demand distribution. Consumer's utility varies depending on the preferences over the product bundle composition; see section 3.2.

Internet access provision tends to be a horizontally integrated industry and the bundle of internet access with other services is an increasingly common contract offered to internet consumers. The so called quadruple play is a commercial bundle of voice, internet access, mobile communications and television services in a single commercial package. The economic reasoning behind bundling is the existence of economies of scope in the provision of telecommunication services. Bundled services share costs of production making the average total cost of the bundle lower than the sum of the individual costs. However, bundling can be constrained by factors like technology. For example, spectrum unavailability may limit the addition of mobile communications to a bundle since the number of mobile operators is limited by the amount of spectrum allocated.

The y dimension represents how a user values package delay or a delay variation that is important for some internet services and applications. Content is heterogeneous in its sensitivity to packet delivery delay. Some types of services such as voice communications over IP are sensitive to the delay, some others are sensitive to the variability of the delay like online gaming and other like file sharing are not sensitive to delay or delay variability at all. For the latter, the best effort approach of internet routing and package transportation is optimal. Consumers have different preferences when using the internet; some of them prefer services such as e-mail or file sharing whereas others also use online games frequently. Some consumers prefer prioritization because of the services they use, and others may prefer neutral networks to enjoy higher network capacity for services that cannot take advantage of prioritization. Consequently, APs locate their internet access products according to consumer preferences to maximize profit. A model extension enabling APs to charge a different price to the consumer segment that prefer prioritization remains for the development of further research. Although with current technology, prioritization is not qualitatively important for most internet services due to technological developments such as protocols that reduce congestion, internet exchange points and peering techniques that reduce packages losses and latency, this approach could be useful for new future innovative services that may highly depend on prioritization such as automated driving. This new services may increase the amount of consumers that prefer prioritization.

There is another important feature of internet access, the connection speed. A higher speed allows for example for the transmission of higher definition video and enables the connection of increased number of devices, or persons using the internet in a household. We assume that all CPs will offer all possible technically feasible connection speeds, thus, in our model, speed is treated as a homogeneous feature and not a distinguishable characteristic of internet access. A model extension considering the maximum speed limitation of some technologies, thus considering speed as a differentiation feature, remains for the development of further future research.

We consider three possible internet access scenarios. Under network neutrality, regulation forces APs to offer standard quality of connection. With network diversity two possibilities are analyzed: a situation where both APs offer prioritized traffic and a scenario where one provider offer prioritized traffic and the other continue under the network neutrality paradigm.

One reason for the co-existence of networks with prioritized and non-prioritized traffic is the presence of economies of density in the deployment of the access network infrastructure. For example, in rural areas there may exist only one fixed network due to the high cost of deployment of a second one. In this situation a fixed broadband provider offering prioritized traffic could coexist with a mobile broadband provider offering residential internet access without priority. The reason for the mobile operator to use NN would be the potential network congestion that may arise in the spectrum based access network if there exists prioritization. In a densely populated area the more plausible scenario would be the presence of two fixed broadband providers, both offering prioritized traffic.

Another situation leading to the co-existence of networks with prioritized and non-prioritized traffic is the use of the access network by the incumbent and an entrant competitor under wholesale access network regulation. To this day, there is no regulation of prioritization on a mandated shared network so the incumbent might be able to offer prioritized traffic whereas the entrant would not. Figures 3.3, 3.4 and 3.5 offer a representation of CP's location and the consumer's indifference curve for the three scenarios analyzed.

There are two APs, AP A and AP B, located at different points of the square of consumer preferences. The location of APs in the square is fixed. Under network neutrality, AP A is located in one of the corners of the square (0, 1) meaning minimum number of product options and non-prioritized traffic and AP B is located in the corner (1, 1) maximum product options and prioritized traffic. Under network diversity there are two possibilities that can be observed in table 3.2.

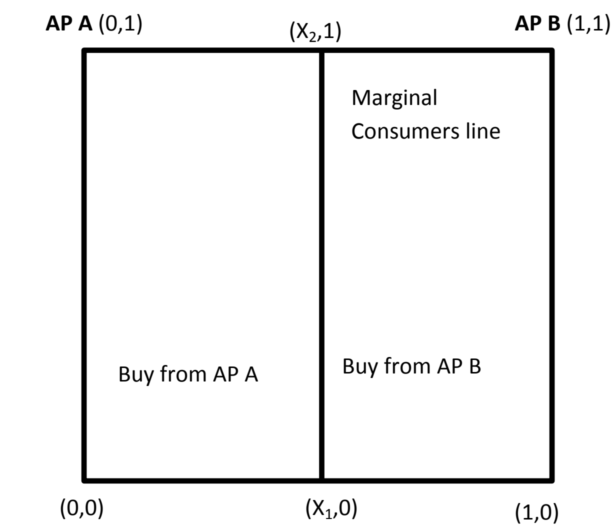
Table 3.6: Access provider's location matrix

	$\mathbf{L_{1A}}$	$\mathbf{L_{2A}}$	$\mathbf{L_{1B}}$	$\mathbf{L_{2B}}$
NN	0	1	1	1
$\mathbf{ND_{BOTH}}$	0	0	1	0
$\mathbf{ND_{ONE}}$	0	1	1	0

Source: Authors

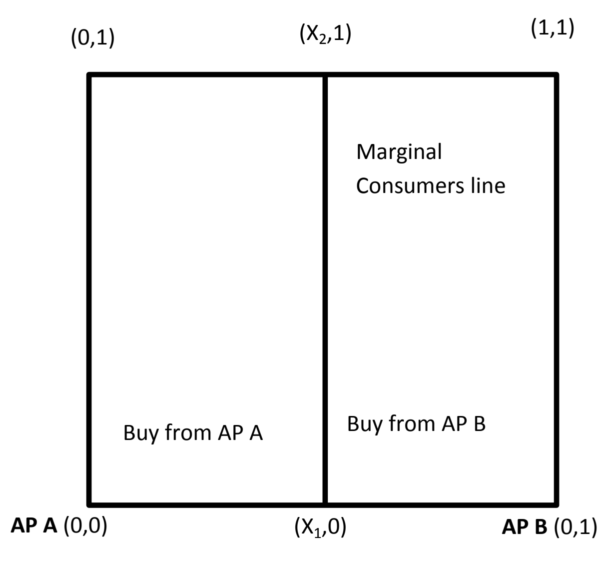
The line passing through $(x_1, 0)$ and $(x_2, 1)$ determines where consumers indifferent from buying internet access from AP A or AP B are located. This line represents the equilibrium in the internet access market, see section 4.1. Consumers to the left of this line buy internet access from AP A. Consumers to the right of the line buy internet from provider B.

Figure 3.10: Access provider's market share under NN



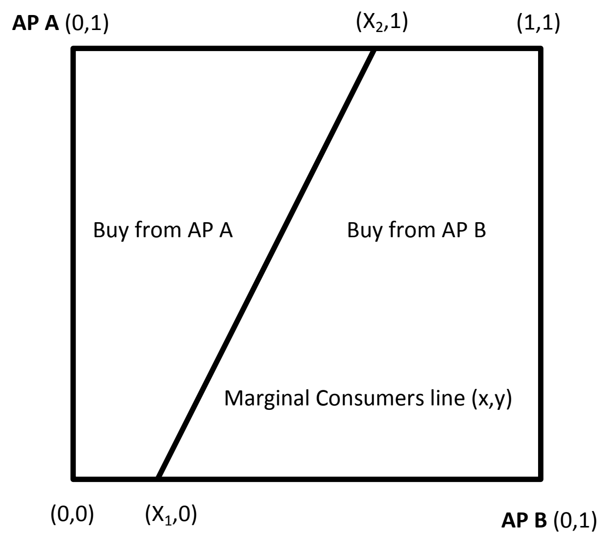
Source: Authors based on Liu and Shuai (2015)

Figure 3.11: Access provider's market share under ND, both with prioritized traffic



Source: Authors based on Liu and Shuai (2015)

Figure 3.12: Access providers' market share under ND with prioritized and non-prioritized traffic



Source: Authors based on Liu and Shuai (2015)

3.2 Internet consumer's utility

Consumers are interested in different services and applications such as web browsing, online purchases, file transfers, video on demand or voice communications and choose an AP according to their tastes. Consumers are heterogeneous in their preferences of internet usage in two dimensions x -preference over products bundled to internet access- and y -network quality in terms of reduced delay of data packages-. Consequently, APs differentiate among them by offering diverse internet access quality options and several related products bundled to internet access.

The model captures consumer's taste heterogeneity defining their utility using a two-dimensional hotelling framework; see Hotelling (1929). We provide an extended version of the one-dimensional model explaining consumer's utility with multi-dimensional product features. Models describing products with multi-dimensional characteristics can be found in these extant studies; Tabuchi (1994), Irmen and Thisse (1998) and Liu and Shuai (2015)).

Consumers are distributed according to their preferences in any position (x, y) of a square of length L x L according to two different distribution functions $F_1\ (x), F_2\ (y)$ where L is normalized to unity for simplicity. The sub-indexes 1 and 2 correspond to the first and second product differentiation dimensions. We use the notation $f_1(x)$, $f_2(y)$ to describe the corresponding density functions.

The utility of a consumer connected to AP A and AP B is defined by:

$$U_A = v\alpha - d(x - \mathrm{L}_{1A})^2 - t\beta(y - \mathrm{L}_{2A})^2 - p_A \quad (1)$$

$$U_B = v\alpha - d(\mathrm{L}_{1B} - x)^2 - t\beta(\mathrm{L}_{2B}(w) - y)^2 - p_B \quad (2)$$

Where the parameters α and β are defined as follows

$$\alpha = \frac{C_p^* + C_{np}^*}{C_p^{pc} + C_{np}^{pc}} \quad (3) \qquad \beta = \frac{C_p^*}{C_p^* + C_{np}^*} \quad (4)$$

U_A and U_B represent the utility enjoyed by internet users when connecting to AP A and AP B respectively, v is the intrinsic value that a consumer obtains from connecting to the internet, α captures the fact that the more content available the higher the utility. It is defined as the ratio of content generated by all CPs under the analyzed scenario to the maximum amount of content that would have been generated under perfect competition. C_p^{pc} and C_{np}^{pc} represent content susceptible of being prioritized and non-prioritized content under perfect competition. C_p^* and C_{np}^* are the prioritized and non-prioritized market equilibrium quantity of content provision under the analyzed regulatory regime.

We assume the value of the product $v\alpha$ to be sufficiently large to cover all costs thus fulfilling the covered market condition. The parameter d is the marginal disutility that a consumers places on the internet product's options when it doesn't match their exact preferences, t is the marginal disutility that a consumers places on the waiting time, β captures the fact that the disutility of delay depends on the availability of prioritized content. It is defined as the proportion of prioritized content to total content in the market equilibrium[14]. The parameters α and β change its value under the different scenarios if content production varies. The terms L_{1A},L_{2A}, L_{1B} and L_{2B} describe the location of AP A and AP B in the square and x and y are the variables representing all possible positions of consumers and APs in the square. The terms w and $\overline{w}$ are the waiting time in the prioritized and non-prioritized networks

For simplicity we normalize the positions of APs with regard to delay in such a way that the only possible values that can be taken in the y axis are 0 or 1, meaning that 1 is the waiting time under network neutrality and 0 is the waiting time in a prioritized network. Thus:

$$\mathrm{L}_{2A}(w) = \mathrm{L}_{2B}(w) = 1 \quad (5) \qquad \mathrm{L}_{2A}(\overline{w}) = \mathrm{L}_{2B}(\overline{w}) = 0 \quad (6)$$

3.3 Impact of a change in consumer preferences on market share

We are looking for a Bertrand's pure strategy Nash equilibrium in prices where firms choose prices simultaneously. A sufficient condition for the existence of this PSNE is that the consumer distribution must be ϱ-concave as defined in Caplin and Nalebuff (1991).

Assumption 1 If the probability density functions $f_1(x)\ f_2(y)$ are ϱ-concave there always exists a PSNE in prices for any value of (d, t) and firm's location.

To determine each AP's market share at equilibrium we identify the line where consumers obtain the same utility from connecting to any of the two existing APs:

$$U_A = U_B \Rightarrow$$

$$v\alpha - d(x - \mathrm{L}_{1A})^2 - t\beta(y - \mathrm{L}_{2A})^2 - p_A = v\alpha - d(\mathrm{L}_{1B} - x)^2 - t\beta(\mathrm{L}_{2B} - y)^2 - p_B$$

For $d > t\beta$ marginal consumers indifferent to connect to AP A or AP B are located on the line:

$$y = \frac{(p_B - p_A) + d(L_{1B}^2 - L_{1A}^2) + t\beta\,(L_{2B}^2 - L_{2A}^2)}{2t\beta(\mathrm{L}_{2B} - \mathrm{L}_{2A})} + \frac{d\,(\mathrm{L}_{1A} - \mathrm{L}_{1B})}{t\beta\,(\mathrm{L}_{2B} - \mathrm{L}_{2A})}x \qquad (7)$$

Consumers to the left of this line buy internet access from AP A. Consumers to the right of the line buy internet from provider B. See figures 3.3, 3.4 and 3.5. Marginal consumers touching the horizontal lines of the square are located on the points $(x_1, 1)$ and $(x_2, 0)$ defined by the following expressions:

$$x_1 = \frac{d\,(L_{1A}^2 - L_{1B}^2) + t\beta\,(L_{2A}^2 - L_{2B}^2) + (p_A - p_B)}{2d(\mathrm{L}_{1A} - \mathrm{L}_{1B})} \qquad (8)$$

$$x_2 = \frac{2t(\mathrm{L}_{2B} - \mathrm{L}_{2A}) + d\,(L_{1A}^2 - L_{1B}^2) + t\beta(L_{2A}^2 - L_{2B}^2) + (p_A - p_B)}{2d(\mathrm{L}_{1A} - \mathrm{L}_{1B})} \qquad (9)$$

The demand function of each AP is given by the following expressions:

$$q_A = F_1(x_1) + \int_{x_1}^{x_2} (1 - F_2(y(x)))\, f_1(x)dx \qquad (10) \qquad\qquad q_B = 1 - q_A \qquad (11)$$

Where $F_1\,(x_1)$ and $F_2\,(y)$ are the distribution functions of consumers in x and y, $f_1(x)$ is the density function of consumers in x and q_A and q_B are the market share of each AP.

A variation in d produces two different effects on the market share of each AP. A rotating effect, that also depends on $t\beta$, and a shifting effect, see figure 3.6. The rotating effect happens because the slope of the indifferent consumers' line depends on d and $t\beta$ in such a way that a decrease in d or an increase in t or in β produces a clockwise rotation of the indifferent consumer line. The shifting effect happens because the magnitude of a shift resulting from a price change in p_A or p_B depends on d. When prices remain unchanged only the rotating effect materializes.

The magnitude of a shift is determined by the following expression:

$$\frac{\partial x_1}{\partial p_A} = \frac{\partial x_2}{\partial p_A} = \frac{1}{2d(\mathrm{L}_{1A} - \mathrm{L}_{1B})} \quad (12)$$

$$\frac{\partial x_1}{\partial p_B} = \frac{\partial x_2}{\partial p_B} = -\frac{1}{2d(\mathrm{L}_{1A} - \mathrm{L}_{1B})} \quad (13)$$

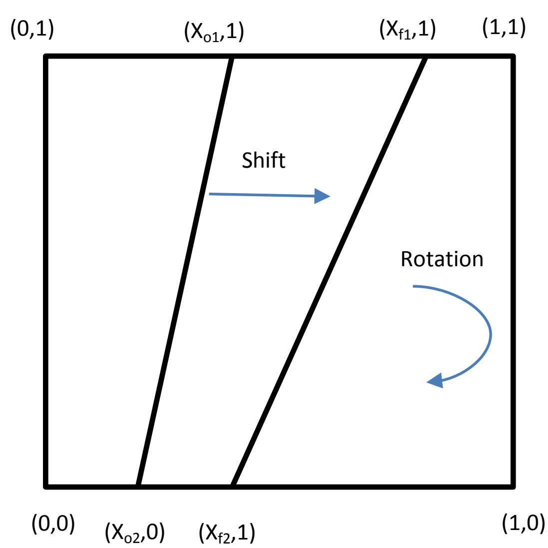
The variation of AP A's market share when there is a shift is represented by:

$$\frac{\partial q_A}{\partial p_A} = f_1(x_1)\frac{\partial x_1}{\partial p_A} + \int_{x_1}^{x_2} \frac{\partial F_2(y)}{\partial p_A} f_1(x)dx \quad (14)$$

The magnitude of the rotating effect is determined by:

$$\frac{\partial y}{\partial x} = \frac{d\ (\mathrm{L}_{1A} - \mathrm{L}_{1B})}{t\beta\ (\mathrm{L}_{2B} - \mathrm{L}_{2A})} \quad (15)$$

Figure 3.13: Shifting and rotating effect of a change in d,t



Source: Authors based on Liu and Shuai (2015)

It is important to note that if the marginal disutility that consumers place on product options d and waiting time t remain constant, a change in the regulatory regime might pose a rotation of the indifferent consumer's line via β, meaning that a change in the quantity of content produced among regulatory regimes lead to different market structures in terms of consumers choosing AP A or AP B.

3.4 Content Provider's market definition and profit maximization problem

CPs obtain revenue from the internet advertising and marketing industry. Advertisers want their ads to be published embedded in relevant content in order to increase sales. Content is more relevant for advertisers the more variety if offers, the more users can potentially consume the content and the more effective is the content to actually increase sales quota. Advertisers pay CPs a price per unit of content c (higher variety), number of internet users that can connect to the CP[15] q and content´s relevance r. The variable c measures the number of data packages per user required to provide the service, q the number of potential consumers and r content's relevance. Content's relevance has to components, the relevance for the internet consumer, r_i^c which measures the effective proportion of users that actually consume the content, and the relevance for the advertiser, r_i^a which is the ability of a content to produce a 'click' to the advertisement that materializes in sales[16]. The value of r is different for each CP.

$$r_i = r_i^c \, r_i^a \tag{16}$$

Content creation has only fixed costs that we normalize to be zero without loss of generality. Thus under network neutrality CPs are homogeneous in terms of costs. In the discriminatory network, however CPs are heterogeneous in terms of costs because of the imposed liability for prioritization charged by APs to those CPs that can increase quality through traffic prioritization. There are a number of CPs N^{cp} of which n_p^{cp} will be forced to pay a fee for traffic prioritization and a number n_{np}^{cp} that will not be forced to pay because service quality do not depend on delay or delay variability.

$$N^{cp} = n_p^{cp} + n_{np}^{cp} \tag{17}$$

CPs are independent monopolist in their own market; consequently they do not compete with each other. We assume inverse lineal demand for content from advertisers where the intercept of each individual demand is different meaning that different services have different data requirements. The factor of relevance r that is also different for each content provider.

$$p(c,r,q) = r_i \, q \, p^a(c) = r_i \, q \, (a_i - bc) \tag{18}$$

Where $p(c,r,q)$ is the price of advertising, $p^a(c)$ is the price of advertising per unit of content, c is the quantity of content generated by a CP per user and a and b are the slope and the intercept of the individual content demand line of each content provider. Internet services have different requirements in terms of data packages to provide the service; therefore a vary for each CP.

In prioritized networks there is a fee per unit of content per user charged by the AP to the content provider that also depends on r_i . We assume that the function is a straight line whose slope is $r_i q$, T can be interpreted as the marginal fee per unit of content.

$$g(c) = T\, r_i q\, c \tag{19}$$

In situations where $g(c)$ is the function describing the fee charged to CPs per unit of content per user from the AP in case of a prioritized network.

Under Network Neutrality each CP's profit can be described by the following expression.

$$\pi^{cp} = r_i(q_A + q_B)\, p^a(c)c \tag{20}$$

Under network Diversity where both APs impose priority CP's profit can be described by the following expressions.

$$\pi_p^{cp} = r_i(q_A + q_B)\ (p^a(c)c - g(c)) \tag{21}$$

$$\pi_{np}^{cp} = r_i(q_A + q_B)\, p^a(c)c \tag{22}$$

Where π_p^{cp} represent the profit of a content provider paying for prioritization and π_{np}^{cp} is the profit of a content provider that is not paying.

Under network diversity where one AP impose priority the other continue under NN, CP's profit can be described by the following expressions

$$\begin{cases} \pi_p^{cp} = r_i((q_A + q_B)p^a(c)c - g(c)\ q_B) & (23) \\ \pi_{np}^{cp} = r_i(q_A + q_B)\, p^a(c)\, c & (24) \end{cases}$$

Content provider's maximization problem for prioritized and non-prioritized content is defined as follows:

$$\underbrace{max}_{c}\, \pi_p^{cp} = r_i(p^a(c)c\ (q_A + q_B) - g(c)q_j) \tag{25}$$

$$\underbrace{max}_{c}\, \pi_{np}^{cp} = r_i(p^a(c)c\ (q_A + q_B)) \tag{26}$$

$$\text{Where} \quad q_j \begin{cases} 0 & with\ NN \\ (q_A + q_B) & with\ ND\ both\ prioritizing \\ q_B & with\ ND\ AP\ B\ prioritizing \end{cases}$$

3.5 Access provider's market definition and profit maximization problem

We assume duopolistic provision of internet access. This assumption is underpinned in the existence of high sunk costs in network deployment. APs obtain revenue from consumers that pay a price for accessing the internet. With network diversity, there are two sources of revenue, the internet access price and a fee obtained from CPs whose services need prioritization.

Under Network Neutrality APs profit only depends on the number of consumers connected:

$$\pi_A = p_A\, q_A \tag{27}$$
$$\pi_B = p_B q_B \tag{28}$$

Under network diversity if both AP prioritize traffic, profit is defined by:

$$\pi_A = (p_A + C_p^* T)\, q_A \tag{29}$$
$$\pi_B = (p_B + C_p^* T)\, q_B \tag{30}$$

Under network diversity with one AP imposing priority and the other continuing under NN paradigm profit is defined by:

$$\pi_A = p_A q_A \tag{31}$$

$$\pi_B = (p_B + C_p^* T)\, q_B \tag{32}$$

Where C_p^* is the sum of content generated by all content providers paying for prioritization per user.

In our sequential game, APs solve two maximization problems; first they maximize their profit with respect to the consumer's internet connection price. Subsequently APs maximize profit with respect to the fee charged to CPs, once CPs have selected their optimal quantity of content to produce. In case of a network with traffic prioritization, maximization problems can be written as follows:

$$\underbrace{max}_{p_j} \pi_j = (p_j + n_p^{cp} g(c^*;T)\ )\, q_j(x,y)\ \ j\epsilon\ [A,B] \tag{33}$$

$$\underbrace{max}_{T} \pi_j(q_A{}^*, q_B{}^*, c^*; T) = (p_j + n_p^{cp} g(c^*;T)\,)\, q_j(x,y)\ \ j\epsilon\ [A,B] \tag{34}$$

In case of a neutral network APs only maximize profit with respect to the price of connection.

$$\underbrace{max}_{p_j} \pi_j = p_j q_j(x,y)\ \ j\epsilon\ [A,B] \tag{35}$$

3.6 Internet advertiser's market definition and consumer surplus

Advertising agencies finance the elaboration of digital content and services consumed by internet users in order to promote the sale of different products. Advertisers obtain a 'consumer surplus' cs from the market of digital services production equal to the difference between what they are willing to pay per unit of content distributed to internet users and how much they actually pay. APs are able to price discriminate services valuing prioritization from other services if network diversity is allowed. In our model, APs are able to extract all the benefits obtained from such price discrimination.

Under Network Neutrality the CS of each advertiser per content provider per internet user is defined by the following expression:

$$cs = \frac{(a_i - p^a(c^*))c^* \ (q_A + q_B)}{2} \tag{36}$$

Under network diversity if both AP prioritize traffic, cs for prioritized traffic is described as follows

$$cs_p = \frac{(a_i - p^a(c_p^*;T^*))c_p^* \ (q_A + q_B)}{2} \tag{37}$$

Consumer surplus in a non-prioritized network is equal to

$$cs_{np} = \frac{(a_i - p^a(c_{np}^*))c_p^* \ (q_A + q_B)}{2} \tag{38}$$

Under network diversity with one AP imposing priority and the other continuing under NN, CS from prioritized traffic is described as follows

$$cs = \frac{(a_i - p^a(c_p^*;T^*))c_p^* \ q_B + (a_i - p^a(c_{np}^*))c_{np}^* \ q_A}{2} \tag{39}$$

Consumer Surplus from non-prioritized traffic is equal to

$$cs_{np} = \frac{(a_i - p^a(c_{np}^*))c_p^* \ (q_A + q_B)}{2} \tag{40}$$

3.7 Total welfare

Total welfare is defined by the following expression

$$TW = PS_{AP} + PS_{CP} + CS_{AD} + CS \tag{41}$$

Where TW is the total welfare composed by the producer surplus of APs PS_{AP}, the producer surplus of CPs PS_{CP} , the consumer surplus of advertisers CS_A,and the consumer surplus of internet users, CS.

4 Equilibrium in the internet access, content distribution and internet advertising markets for different scenarios

4.1 Equilibrium price of internet access, equilibrium quantity of content generated and total welfare under network neutrality

Under NN there is no difference in terms of delay in both AP's networks, consequently the location of the two APs is given by $L_{2B} = L_{2A} = 1$. Under NN the only differentiation strategy is to offer services bundled with the internet access different of those of the competition. In our model AP A choose to offer only internet access whereas AP B offer a set of bundled services, thus $L_{2A} = 1\,; L_{2B} = 0$. Consumers obtaining the same utility from connecting to any of the two existing CPs are described by:

$$U_A = U_B \Rightarrow$$

$$v\alpha - dx^2 - t\beta(y - 1)^2 - p_A = v\alpha - d(1-x)^2 - t\beta(1-y)^2 - p_B$$

Marginal consumers indifferent to connect to AP A or AP B are located at the vertical line:

$$x = \frac{(p_B - p_A)}{2d} + \frac{1}{2} \tag{42}$$

Marginal consumers touching the horizontal lines of the square are located at

$$x_1 = \frac{(p_B - p_A)}{2d} + \frac{1}{2} \tag{43}$$

$$x_2 = \frac{(p_B - p_A)}{2d} + \frac{1}{2} \tag{44}$$

As we are looking for a Bertrand PSNE, in the equilibrium $p_A^* = p_B^*$ therefore:

$$x^* \equiv x_1^* = x_2^* = \frac{1}{2}$$

The shifting effect is defined by:

$$\frac{\partial x_1}{\partial p_A} = \frac{\partial x_2}{\partial p_A} = -\frac{1}{2d} \tag{45}$$

$$\frac{\partial x_1}{\partial p_B} = \frac{\partial x_2}{\partial p_B} = \frac{1}{2d} \tag{46}$$

Under NN there is no rotation effect; AP's profit is determined by market share and price of connection

$$\pi_A = p_A\ q_A\ = p_A F_1(x_1) \tag{47}$$

$$\pi_B = p_B q_B = p_B\ (1 - q_A) = p_B\ (1 - F_1(x_1)) = p_B(F_1(x_2) - F_1(x_1)) \tag{48}$$

Proposition 1: In the first stage of the sequential-moves game in the unique pure strategy Bertrand-Nash equilibrium, firms split the market $q_A = q_B = \frac{1}{2}$ and choose a price:

$$p \equiv p_A^* = p_B^* = \frac{d}{f_1\left(\frac{1}{2}\right)}$$

Proof. See appendix

Proposition 2: Equilibrium price varies with a change in internet consumer preferences in the product options dimension d according to the following expression:

$$\frac{\partial p^*}{\partial d} = \frac{1}{f_1\left(\frac{1}{2}\right)}$$

Proof: See appendix

Proposition 3: Under network neutrality, in the second stage of the sequential-moves game, the equilibrium quantity of content per user generated by a CP and the total equilibrium quantity of content per user is equal to:

$$c^* = \frac{a_i}{2b} \qquad , \qquad C^* = \sum_{i=1}^{N^{cp}} \frac{a_i}{2b}$$

Proof. See appendix

Proposition 4: After the two stages of the sequential moves game, total welfare described as the sum of producer surplus of access providers and content providers and the consumer surplus of internet users and advertisers is defined by the expressions below:

The profit obtained by AP A is $\pi_A = \frac{d}{f_1\left(\frac{1}{2}\right)} q_A$, the profit obtained by AP B is $\pi_B = \frac{d}{f_1\left(\frac{1}{2}\right)} q_B$ and $q_A = q_B = \frac{1}{2}$, therefore total producer surplus of access providers is:

$$PS_{AP} = \pi_A + \pi_B = \frac{d}{f_1\left(\frac{1}{2}\right)}$$

The profit obtained by each content provider is $\pi^{cp} = \frac{(q_A + q_B)}{4b} r_i a_i^2$, therefore total producer surplus is equal to:

$$PS^{cp} = \frac{1}{4b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)$$

The consumer surplus of all the advertisers is the sum of the surpluses of all advertisers

$$CS = \frac{1}{8b}\sum_{i=1}^{N^{cp}} r_i a_i^2 \quad \rightarrow CS = \frac{1}{8b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)$$

The consumer surplus of internet consumers is

$$CS_A^{nn} = \int_0^1 \int_0^{x_1^*} f_1(x) f_2(y)\, U_A^{nn}(x,y)\, dxdy$$

$$CS_B^{nn} = \int_0^1 \int_{x_1^*}^1 f_1(x) f_2(y)\, U_A^{nn}(x,y)\, dxdy$$

Total welfare is

$$TW = \frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{3}{4b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right) + CS^{nn}(f_1(x), f_2(x), U^{nn}(x,y); v, \alpha, \beta, t, d)$$

Proof. See appendix

4.2 Equilibrium price of internet access, tax charged to content providers and quantity of content in a discriminatory network where both APs offer prioritized traffic

Under ND where both APs offer prioritized traffic the location of the two APs is given by $\mathrm{L}_{2B} = \mathrm{L}_{2A} = 0$. In the model AP A choose to offer only internet access whereas AP B offer a set of bundled services, thus $\mathrm{L}_{2A} = 0\,; \mathrm{L}_{2B} = 0$. Consumers obtaining the same utility from connecting to any of the two existing CPs are described by:

$$U_A = U_B \Rightarrow$$

$$v\alpha - dx^2 - t\beta y^2 - p_A = v\alpha - d(1-x)^2 - t\beta y^2 - p_B$$

Marginal consumers indifferent to connect to AP A or AP B are located at the vertical line:

$$x = \frac{(p_B - p_A)}{2d} + \frac{1}{2} \tag{49}$$

The result is identical to that obtained in the NN case, therefore marginal consumers touching the horizontal lines of the square are located at $x_1 = \frac{(p_B - p_A)}{2d} + \frac{1}{2}$; $x_2 = \frac{(p_B - p_A)}{2d} + \frac{1}{2}$, $p_A^* = p_B^*$ $x^* \equiv x_1^* = x_2^* = \frac{1}{2}$ and the shifting effect is defined by $\frac{\partial x_1}{\partial p_A} = \frac{\partial x_2}{\partial p_A} = -\frac{1}{2d}$; $\frac{\partial x_1}{\partial p_B} = \frac{\partial x_2}{\partial p_B} = \frac{1}{2d}$. As in the NN case there is no rotation effect.

The expressions for price and the impact of a variation in consumer preferences obtained in propositions 1 and 2 for the NN case are equivalent to the discriminatory network where both APs offer prioritized traffic.

Proposition 5: Under network diversity where both APs offer prioritized traffic the equilibrium tax charged by the AP to the CP is defined by:

$$T^* = \frac{\sum_{i=1}^{n_p^{cp}} a_i}{n_p^{cp}}$$

Therefore equilibrium quantity of prioritized and non-prioritized content generated by a CP is

$$c_p^* = \frac{1}{2b}\left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right] \quad ; \quad c_{np}^* = \frac{a_i}{2b}$$

Total prioritized and non-prioritized content per internet user is

$$C_p^* = \sum_{i=1}^{n_p^{cp}} \frac{a_i}{4b} \quad ; \quad C_{np}^* = \sum_{i=1}^{n_{np}^{cp}} \frac{a_i}{2b}$$

Total internet traffic in each network is

$$C^* = C_p^* + C_{np}^* = \frac{1}{2b}\left[\sum_{i=1}^{n_p^{cp}} \frac{a_i}{2} + \sum_{i=1}^{n_{np}^{cp}} a_i\right]$$

Proposition 6: After the two stages of the sequential moves game, total welfare described as the sum of producer surplus of access providers and content providers and the consumer surplus of internet users and advertisers is defined by the expressions below:

The profit obtained by CPs prioritizing and non-prioritizing traffic is $\pi_p^{cp} = \frac{r_i(q_A+q_B)}{4b}\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2$ and $\pi_{np}^{cp} = \frac{(q_A+q_B)}{4b} r_i a_i^2$. The total profit for all CPs, producer surplus in the content provision market is:

$$PS^{cp} = \frac{1}{4b}\left[\sum_{i=1}^{n_p^{cp}} r_i\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right]$$

The producer surplus defined as the sum of profit of AP A and AP B is

$$PS_{AP} = \frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{1}{8bn_p^{cp}}\sum_{i=1}^{n_p^{cp}} a_i^2$$

The consumer surplus of all the advertisers is the sum of the surpluses of advertisers that pay for prioritization $CS_p = \frac{1}{16b}\left[\sum_{i=1}^{n_p^{cp}} r_i(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}})^2\right]$ and those that do not pay $CS_{np} = \frac{(q_A+q_B)}{8b}\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2$

$$CS^a = CS_p + CS_{np} = \frac{1}{16b}\left[\sum_{i=1}^{n_p^{cp}} r_i\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right]$$

The consumer surplus of internet users is

$$CS_A^{ndb} = \int_0^1 \int_0^{x_1^*} f_1(x) f_2(y)\, U_A^{ndb}(x,y)\, dxdy$$

$$CS_B^{ndb} = \int_0^1 \int_{x_1^*}^1 f_1(x) f_2(y)\, U_A^{ndb}(x,y)\, dxdy$$

Total welfare is

$$TW = \frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{5}{16b}\left[\sum_{i=1}^{n_p^{cp}} r_i\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right] + CS^{ndb}$$

Proof. See appendix

4.3 Equilibrium price of internet access, tax charged to content providers and quantity of content in a discriminatory network where one AP is offering prioritized traffic and the other is acting as under network neutrality regulation

In this case APs can differentiate in the two differentiation dimensions: network quality and bundling options. We assume one AP will choose to create traffic classes whereas the other will continue using the best effort approach. In this situation $\overline{w} \neq w$ and $\mathrm{L}_{2A} = 1\,; \mathrm{L}_{2B} = 0$. In our model AP A choose to offer only internet access whereas AP B offer bundled services, thus $\mathrm{L}_{2A} = 1\,; \mathrm{L}_{2B} = 0$. Consumers obtaining the same utility from connecting to any of the two existing CPs are described by:

$$v\alpha - d(x - \mathrm{L}_{1A})^2 - t\beta(y - \mathrm{L}_{2A})^2 - p_A = v\alpha - d(\mathrm{L}_{1B} - x)^2 - t\beta(\mathrm{L}_{2B} - y)^2 - p_B$$

$$\widetilde{U_A} = \widetilde{U_B} \Rightarrow$$

$$v\alpha - dx^2 - t\beta(y-1)^2 - p_A = v\alpha - d(1-x)^2 - t\beta y^2 - p_B$$

Marginal consumers indifferent to connect to AP A or AP B are located at the following line

$$\tilde{y} = \frac{(p_B - p_A) + (d - t\beta)}{2t\beta} + \frac{d}{t\beta}x \qquad (50)$$

Marginal consumers touching the horizontal lines of the square are located at

$$\widetilde{x_1} = \frac{(p_B - p_A) + (d - t\beta)}{2d} \qquad (51) \qquad\qquad \widetilde{x_2} = \frac{(p_B - p_A) + (d + t\beta)}{2d} \qquad (52)$$

As we are looking for a Bertrand PSNE, in the equilibrium $p_A = p_B$ therefore

$$\tilde{y} = \frac{(d-t)}{2t} - \frac{d}{t\beta}x$$

$$\widetilde{x_1^*} = \frac{d+t}{2d} = \frac{1}{2} + \frac{t\beta}{2d} \qquad\qquad \widetilde{x_2^*} = \frac{d - t\beta}{2d} = \frac{1}{2} - \frac{t\beta}{2d}$$

The shifting effect is defined by:

$$\frac{\partial \widetilde{x_1}}{\partial p_A} = \frac{\partial \widetilde{x_2}}{\partial p_A} = -\frac{1}{2d} \qquad (53) \qquad\qquad \frac{\partial \widetilde{x_1}}{\partial P_B} = \frac{\partial \widetilde{x_2}}{\partial P_B} = \frac{1}{2d} \qquad (54)$$

And the rotating effect by:

$$\frac{\partial y}{\partial x} = \frac{d}{t\beta} \tag{55}$$

Proposition 7: Under network diversity when one firm prioritize and the other doesn´t and firms are located in the upper left and lower right of the square in the unique PSBNE firms split the market $q_A = q_B = \frac{1}{2}$ and choose a price:

$$p^* \equiv p_A^* = p_B^* = \frac{d}{\int_{x_1}^{x_2} f_2(y) f_1(x) dx}$$

Proof: See appendix

Proposition 8: Under network diversity when one firm prioritizes and the other doesn't, equilibrium price may increase or decrease with a change in consumer preferences in either product differentiation dimension depending on the signs of the partial derivatives shown as follows.

$$\frac{\partial p^*}{\partial d} = \frac{\int_{x_1^*}^{x_2^*} f_2(y) f_1(x) dx - \left(-f_2(y\left(\frac{d}{t\beta}\right)) f_1(x_2^*) \frac{t\beta}{2d^2} - f_2\left(y(\frac{d}{t\beta} - 1)\right) f_1(x_1^*) \frac{t\beta}{2d^2}\right) + \int_{x_2^*}^{x_2^*} f_2'(y) \left(\frac{1}{2t\beta} - \frac{1}{t\beta} x\right) f_1(x) dx}{(\int_{x_1^*}^{x_2^*} f_2(y) f_1(x) dx)^2}$$

$$\frac{\partial p^*}{\partial t} = \frac{\int_{x_1^*}^{x_2^*} f_2(y) f_1(x) dx - \left(f_2(y\left(\frac{d}{t\beta}\right)) f_1(x_2^*) \frac{\beta}{2d} + f_2\left(y(\frac{d}{t\beta} - 1)\right) f_1(x_1^*) \frac{\beta}{2d}\right) + \int_{x_2^*}^{x_2^*} f_2'(y) \left(\frac{(t - 2t\beta)\beta - d}{2t^2\beta^2} - \frac{d}{t^2\beta} x\right) f_1(x) dx}{(\int_{x_1^*}^{x_2^*} f_2(y) f_1(x) dx)^2}$$

Proof: See appendix

Proposition 9: Under network diversity where one APs offer prioritized traffic and the other operates under NN, the equilibrium tax charged by the AP to the CP in the discriminatory network is defined by:

$$T^* = \frac{(q_A + q_B) \sum_{i=1}^{n_p^{cp}} a_i}{2 n_p^{cp} q_B}$$

Prioritized and non-prioritized content per content provider and internet user is

$$c_p^* = \frac{1}{2b}\left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right] \qquad c_{np}^* = \frac{a_i}{2b}$$

Total prioritized and non-prioritized content per internet user is

$$C_p^* = \sum_{i=1}^{n_{np}^{cp}} \frac{a_i}{4b} \qquad C_{np}^* = \sum_{i=1}^{n_{np}^{cp}} \frac{a_i}{2b}$$

Total traffic in all networks is

$$C^* = C_p^* + C_{np}^* = \frac{1}{2b}\left[\sum_{i=1}^{n_p^{cp}} \frac{a_i}{2} + \sum_{i=1}^{n_{np}^{cp}} a_i\right]$$

Proof: See appendix

Proposition 10: After the two stages of the sequential moves game, total welfare described as the sum of producer surplus of access providers and content providers and the consumer surplus of internet users and advertisers is defined by the expressions below:

The producer surplus of access providers is

$$PS_{AP} = \pi_A + \pi_B = \frac{d}{\int_{x_1}^{x_2} f_2(y) f_1(x) dx)} + \frac{1}{8bn_p^{cp}} \sum_{i=1}^{n_p^{cp}} a_i^2$$

The producer surplus of content providers is

$$PS^{cp} = \frac{1}{4b}\left(\frac{1}{2}\sum_{i=1}^{n_p^{cp}} r_i {a_i}^2 + \sum_{i=1}^{n_{np}^{cp}} r_i {a_i}^2\right)$$

The consumer surplus of advertisers is

$$CS = \frac{1}{16b}\left[\left(\sum_{i=1}^{n_p^{cp}} r_i \left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]^2 + 4\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)\right]$$

Internet user's consumer surplus is

$$CS^{ndo} = CS_A^{ndo} + CS_B^{ndo}$$

$$CS_A^{ndo} = \int_0^1 \int_0^{x_1^*} f_1(x) f_2(y)\, U_A^{ndo}(x,y)\, dxdy + \int_0^1 \int_{x_1^*}^{x_2^*} (1 - F_2(y(x))\, f_1(x)\, U_A^{ndo}(x,y)\, dxdy$$

$$CS_B^{ndo} = \int_0^1 \int_0^{x_1^*} F_2(y(x)) f_1(x)\, U_A^{ndo}(x,y)\, dxdy + \int_0^1 \int_{x_1^*}^{x_2^*} f_1(x) f_2(y)\, U_A^{ndo}(x,y)\, dxdy$$

Total Welfare is

$$TW = \frac{d}{\int_{x_1}^{x_2} f_2(x) f_1(x) dx)} + \frac{1}{8bn_p^{cp}} \sum_{i=1}^{n_p^{cp}} a_i^2 + \frac{1}{4b}\left(\frac{1}{2}\sum_{i=1}^{n_p^{cp}} r_i {a_i}^2 + \sum_{i=1}^{n_{np}^{cp}} r_i {a_i}^2\right) + \frac{1}{16b}\left[\left(\sum_{i=1}^{n_p^{cp}} r_i \left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]^2 + 4\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)\right] + CS^{ndo}$$

Proof. See appendix

5. Simulations

In this section we offer systematic simulations using feasible ranges for parameters value to help discern the impact of departing from network neutrality regulation on social welfare. The objective of the simulations is to fill the gaps that cannot be easily explained using only mathematical analysis. A summary of the main parameters value used in the simulations can be consulted in annex 2.

5.1 Methodology

The model considers three different broadband access price discrimination scenarios. Under network neutrality, regulation forces APs to provide treat all data packages equally and uniform pricing. With network diversity two options are analyzed. One option with both APs offering prioritized traffic and a case where one AP offers prioritized traffic and the other standard quality. The two latter cases might coexist at the same time in different geographical areas. For example, in densely populated areas both APs may be willing to charge for traffic prioritization whereas in rural locations prioritization may be suitable only for one AP.

Simulations are carried out considering two different distributions of consumers on preferences for each price discrimination scenario, uniform and normal distributions. Furthermore, we consider four different combinations of feasible ranges of the relevance r_i and the price of content a_i parameters. The four options are the result of combining concentrated and disperse ranges for each of these two parameters. The values for a_i, r_i have been generated randomly using Montecarlo simulation modelling. Consequently, for each price discrimination scenario we present eight tabulated values, four of them corresponding to the normal distribution function and other four to the uniform distribution.

5.2 Simulation results

We present simulation results in tables showing total welfare and each of its components -producer surplus of access providers, producer surplus of content providers, consumer surplus of internet advertisers and the consumer surplus of internet users-. Results are expressed as a proportion of the consumer's intrinsic value of internet connectivity v which is set to 1. Together with the tables we also present the standard deviation values.

Table 7.3: Total Welfare

TOTAL WELFARE		Concentrated a_i Concentrated r_i	Concentrated a_i Disperse r_i	Disperse a_i Concentrated r_i	Disperse a_i Disperse r_i
Network Neutrality	Uniform	0.5566205	0.6167692	0.5774567	0.6849972
	Normal	0.5228772	0.5831186	0.5437084	0.6512734
Network Diversity Both APs	Uniform	0.4053743	0.4555483	0.4206971	0.5087805
	Normal	0.3830296	0.4344061	0.3983819	0.4873571
Network Diversity One AP	Uniform	0.4127036	0.4547915	0.4278398	0.5088357
	Normal	0.4127169	0.4562172	0.4281580	0.5113532

Source: Authors

Table 3.8: Total welfare standard deviation

TOTAL WELFARE (Standard deviation)		Concentrated a_i Concentrated r_i	Concentrated a_i Disperse r_i	Disperse a_i Concentrated r_i	Disperse a_i Disperse r_i
Network Neutrality	Uniform	0.0052771	0.0093293	0.0103178	0.0239929
	Normal	0.0052520	0.0094216	0.0103491	0.0241024
Network Diversity Both APs	Uniform	0.0059820	0.0158290	0.0060209	0.0167423
	Normal	0.0057773	0.0145019	0.0059049	0.0165380
Network Diversity One AP	Uniform	0.0059352	0.0140984	0.0065157	0.0165960
	Normal	0.0071922	0.0184738	0.0073156	0.0183097

Source: Authors

Table 3.9: Producer surplus of Access providers

PRODUCER SURPLUS OF APs		Concentrated a_i Concentrated r_i	Concentrated a_i Disperse r_i	Disperse a_i Concentrated r_i	Disperse a_i Disperse r_i
Network Neutrality	Uniform	0.0600000	0.0600000	0.0600000	0.0600000
	Normal	0.0376295	0.0376295	0.0376295	0.0376295
Network Diversity Both APs	Uniform	0.0771816	0.0865538	0.0777343	0.0866442
	Normal	0.0548111	0.0641833	0.0553638	0.0642738
Network Diversity One AP	Uniform	0.3316551	0.3308434	0.3330778	0.3401585
	Normal	0.3316551	0.3308434	0.3330778	0.3401585

Source: Authors

Table 3.10: Standard deviation of the producer surplus of Access providers

PRODUCER SURPLUS OF APs (Standard Deviation)		Concentrated a_i Concentrated r_i	Concentrated a_i Disperse r_i	Disperse a_i Concentrated r_i	Disperse a_i Disperse r_i
Network Neutrality	Uniform	0.0000000	0.0000000	0.0000000	0.0000000
	Normal	0.0000000	0.0000000	0.0000000	0.0000000
Network Diversity Both APs	Uniform	0.0009280	0.0044888	0.0016176	0.0038369
	Normal	0.0009280	0.0044888	0.0016176	0.0038369
Network Diversity One AP	Uniform	0.0087615	0.0362713	0.0115402	0.0305613
	Normal	0.0087615	0.0362713	0.0115402	0.0305613

Source: Authors

Table 3.11: Producer surplus of content providers

PRODUCER SURPLUS OF CPs		Concentrated a_i Concentrated r_i	Concentrated a_i Disperse r_i	Disperse a_i Concentrated r_i	Disperse a_i Disperse r_i
Network Neutrality	Uniform	0.0772727	0.1257934	0.0939196	0.1800584
	Normal	0.0772727	0.1257934	0.0939196	0.1800584
Network Diversity Both APs	Uniform	0.0507983	0.0938380	0.0629000	0.1341425
	Normal	0.0507983	0.0938380	0.0629000	0.1341425
Network Diversity One AP	Uniform	0.0581276	0.0930811	0.0700427	0.1341978
	Normal	0.0581276	0.0930811	0.0700427	0.1341978

Source: Authors

Table 3.12: Standard deviation of the producer surplus of content providers

PRODUCER SURPLUS OF CPs (Standard Deviation)		Concentrated a_i Concentrated r_i	Concentrated a_i Disperse r_i	Disperse a_i Concentrated r_i	Disperse a_i Disperse r_i
Network Neutrality	Uniform	0.0041189	0.0079409	0.0083963	0.0196852
	Normal	0.0041189	0.0079409	0.0083963	0.0196852
Network Diversity Both APs	Uniform	0.0032900	0.0051562	0.0045561	0.0132643
	Normal	0.0032900	0.0051562	0.0045561	0.0132643
Network Diversity One AP	Uniform	0.0034586	0.0043909	0.0053945	0.0139830
	Normal	0.0034586	0.0043909	0.0053945	0.0139830

Source: Authors

Table 3.13: Consumer surplus, advertisers

CONSUMER SURPLUS OF ADVERTISERS		Concentrated a_i Concentrated r_i	Concentrated a_i Disperse r_i	Disperse a_i Concentrated r_i	Disperse a_i Disperse r_i
Network Neutrality	Uniform	0.0193182	0.0314483	0.0234799	0.0450146
	Normal	0.0193182	0.0314483	0.0234799	0.0450146
Network Diversity Both APs	Uniform	0.0126996	0.0234595	0.0157250	0.0335356
	Normal	0.0126996	0.0234595	0.0157250	0.0335356
Network Diversity One AP	Uniform	0.0126996	0.0234595	0.0157250	0.0335356
	Normal	0.0126996	0.0234595	0.0157250	0.0335356

Source: Authors

Table 3.14: Standard deviation of the consumer surplus of advertisers

CONSUMER SURPLUS OF ADVERTISERS (Standard Deviation)		Concentrated a_i Concentrated r_i	Concentrated a_i Disperse r_i	Disperse a_i Concentrated r_i	Disperse a_i Disperse r_i
Network Neutrality	Uniform	0.0010297	0.0019852	0.0020991	0.0049213
	Normal	0.0010297	0.0019852	0.0020991	0.0049213
Network Diversity Both APs	Uniform	0.0008225	0.0012890	0.0011390	0.0033161
	Normal	0.0008225	0.0012890	0.0011390	0.0033161
Network Diversity One AP	Uniform	0.0008225	0.0012890	0.0011390	0.0033161
	Normal	0.0008225	0.0012890	0.0011390	0.0033161

Source: Authors

Table 3.15: Consumer surplus

CONSUMER SURPLUS OF USERS		Concentrated a_i Concentrated r_i	Concentrated a_i Disperse r_i	Disperse a_i Concentrated r_i	Disperse a_i Disperse r_i
Network Neutrality	Uniform	0.4000296	0.3995275	0.4000573	0.3999242
	Normal	0.3886568	0.3882474	0.3886794	0.3885708
Network Diversity Both APs	Uniform	0.2646948	0.2516970	0.2643378	0.2544581
	Normal	0.2647206	0.2529252	0.2643931	0.2554052
Network Diversity One AP	Uniform	0.0102214	0.0074074	0.0089942	0.0009438
	Normal	0.0102346	0.0088331	0.0093124	0.0034612

Source: Authors

Table 3.16: Standard deviation of the consumer surplus

CONSUMER SURPLUS OF USERS (Standard Deviation)		Concentrated a_i Concentrated r_i	Concentrated a_i Disperse r_i	Disperse a_i Concentrated r_i	Disperse a_i Disperse r_i
Network Neutrality	Uniform	0.0003435	0.0016075	0.0004788	0.0012507
	Normal	0.0002801	0.0013109	0.0003904	0.0010200
Network Diversity Both APs	Uniform	0.0029635	0.0154818	0.0048478	0.0124493
	Normal	0.0026721	0.0139662	0.0043769	0.0112378
Network Diversity One AP	Uniform	0.0063278	0.0248572	0.0081416	0.0216809
	Normal	0.0049127	0.0199791	0.0062969	0.0166440

Source: Authors

6. Conclusions

We present a model to analyze total welfare in two sided markets where there is imperfect competition; the platform is able to differentiate its product in multiple dimensions on one side and to price discriminate on the other side. The model has been customized to analyze the impact of different internet access regulatory regimes on total welfare but with not very significant changes it could be used to undertake other different type of industrial organization and regulatory analysis in markets with similar structures.

The model aims to shed some light on the immediate effect of a regulatory action on total welfare. It is not intended to measure the impact of regulation in the long run. In other words, it doesn´t evaluate the incentives of APs and CPs to invest and innovate in the future. This calculation is left for future research. However, we consider the immediate impact of the regulatory action more suitable since only a percentage of the new available to APs resources may be used for network expansion and innovation. The elaboration of digital content is financed through internet advertising but we can expect similar results for content financed by internet consumers. We leave this analysis for future research.

Conclusions are drawn from simulation results and from the set of propositions described throughout the paper. Proofs can be found on appendix A. Tables comparing propositions for the three different internet access scenarios are provided in Annex 1. Simulation result tables are available in section 5.2, and parameter values used in the simulation can be found in Annex2.

Abandoning network neutrality regulation reduces the quantity of internet content produced according to propositions 6 and 10. Network diversity produces a significant reduction of prioritized content while the quantity of content non-susceptible of being prioritized remains unchanged. The reduction in the quantity of content is equal in the two network diversity scenarios – both and only one AP prioritizing - thus, even when only one AP offers prioritized content the reduction is materialized on equal footing.

According to propositions 4 and 6, under network diversity APs enjoy higher income coming from the tax charged to CPs for prioritization but CPs offering prioritized content, and advertisers suffer an income reduction. Income reduction depends on the sum of prices paid by advertisers a_i to CPs and also on consumer preferences when only one AP offer prioritization. Firms supplying non-prioritized content enjoy the same surplus. Abandoning network neutrality has two different types of effects on consumer's utility. On one hand, it reduces utility through a decrease in the value of α –ratio of content actually available to content available with perfect competition-, on the other hand, the disutility caused by the delay of data packages is reduced via β –ratio of prioritized content to content not susceptible of prioritization -. The simulation is necessary to measure net impact.

According to propositions 2 and 8, a change in the regulatory regime would lead to different internet access market structures because there is a rotation of the indifferent consumer´s line via changes in β. Even when the marginal disutility that consumers place on product options d and waiting time t remains constant, the change in β would rotate the indifferent consumer´s line.

Results are expressed with respect to v which is set to 1. v is the maximum utility that a consumer could obtain with perfect competition in the provision of internet content thus with maximum availability of content. However, consumers usually don't enjoy v but $v\alpha$ where the multiplier α lowers utility when there is reduced competition and/or a tax on content production.

Simulations show that network neutrality regulation is welfare superior to network diversity under the model assumptions with the values and value ranges given to parameters. Departing from network neutrality regulation leads to an abrupt decrease of the consumer surplus of internet users, from 0.4 to between 0.01 and 0.26. This effect weights much more in total welfare than the increase in APs surplus coming from the tax charged to content providers.

The increase in the producer surplus of access providers, from 0.06 to a maximum of 0.34 is not enough to compensate the reduction of the consumer surplus of internet users. Among the two network diversity scenarios, results are very similar with slightly higher total welfare results in the case of only one CP offering prioritized traffic. Disperse range of parameter values for r_i and a_i tend to produce higher total welfare.

We can set a monetary value for v to make simulation results more intuitive. For example, if we set v to 150$ per month the simulation shows that total welfare would range from 57$ to 102$, consumer's utility between 1$ to 60$, and average ARPU for APs between 9$ to 51$, for CPs between 8$ to 20$, and for advertisers between 2$ to 6.8$. Although, the values for a_i, r_i have been generated randomly using Montecarlo simulation modelling, they have a significant relation with real observed ARPU values[17].

Appendix A: Proof of Propositions

Proof of Proposition 1

AP A's profit with network neutrality is equal to $\pi_A = p_A\, q_A\ = p_A F_1(x_1)$, where $x_1 = \frac{(p_B - p_A)}{2d} + \frac{1}{2}$ and $\frac{\partial x_1}{\partial p_A} = -\frac{1}{2d}$, taking the derivative of π_A with respect to p_A we obtain the price in the equilibrium p_A^*. As we are looking for the Pure Strategy Bertrand-Nash equilibrium we know that: $p_A^* = p_B^*$, $x_1^* = x_2^* = \frac{1}{2}$ and $F_1(x_1^*) = \frac{1}{2}$ therefore:

$$\frac{\partial \pi_A}{\partial p_A} = F_1(x_1^*) + p_A f_1(x_1^*)\frac{\partial x_1}{\partial p_A} = 0 \rightarrow F_1(x_1^*) - \frac{p_A f_1(x_1^*)}{2d} = 0$$

$$p_A^* = p_B^* = \frac{2dF_1(x_1)}{f_1(x_1^*)} = \frac{d}{f_1(x_1^*)} = \frac{d}{f_1\left(\frac{1}{2}\right)}$$

Proof of Proposition 2

Using the results from proposition 1 $p_A^* = p_B^* = \frac{d}{f_1\left(\frac{1}{2}\right)}$ taking the derivative of p_A^* with respect to d we obtain the effect of a change in d on the equilibrium price:

$$\frac{\partial p_A^*}{\partial d} = \frac{1}{f_1\left(\frac{1}{2}\right)}$$

Proof of Proposition 3

We defined each CP's profit with the following expression $\pi^{cp} = p^a(c)c\ (q_A + q_B)r_i = (a_i - bc)c\ (q_A + q_B)r_i$ Taking the derivative of π^{cp} with respect to c and making this expression equal to zero (F.O.C.) we can calculate the equilibrium quantity per CP:

$$\frac{\partial \pi_p^{cp}}{\partial c} = r_i(q_A + q_B)\,(a_i - 2bc) = 0 \ \rightarrow\ c^* = \frac{a_i}{2b}$$

The total content per internet user is the sum of the content transmitted by all CP

$$C^* = \sum_{i=1}^{N^{cp}} \frac{a_i}{2b}$$

Proof of Proposition 4

From proposition 3 we know that $c^* = \frac{a_i}{2b}$ therefore we can calculate each CP profit as an expression of a_i and b

$$\pi^{cp} = r_i\, p^a(c^*)c^*\ (q_A + q_B) = r_i(q_A + q_B)(a_i c^* - b(c^*)^2) = r_i(q_A + q_B)\left(\frac{{a_i}^2}{2b} - \frac{a_i^2}{4b}\right)$$

$$\pi^{cp} = \frac{(q_A + q_B)}{4b} r_i a_i^2$$

The sum of the profit of all CPs is the producer surplus in the content provision market

$$PS^{cp} = \frac{(q_A + q_B)}{4b} \sum_{i=1}^{N^{cp}} r_i a_i^2$$

Considering that $q_A = q_B = \frac{1}{2}$

$$PS^{cp} = \frac{1}{4b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)$$

Under Network Neutrality the profit of APs is defined by $\pi_A = p_A\, q_A$ and $\pi_B = p_B q_B$ and $p_A^* = p_B^* = \frac{d}{f_1\left(\frac{1}{2}\right)}$ and $q_A = q_B = \frac{1}{2}$ then

$$PS_{AP} = \pi_A + \pi_B = \frac{d}{f_1\left(\frac{1}{2}\right)}$$

We defined the consumer surplus of each advertiser per unit of content per user under NN as $cs = \frac{r_i(a_i - p^a(c^*))c^*\ (q_A+q_B)}{2}$ where $c^* = \frac{a_i}{2b}$ and $q_A = q_B = \frac{1}{2}$ therefore:

$$cs = \frac{r_i(a_i - (a_i - bc^*))c^*\ (q_A + q_B)}{2} = \frac{r_i(q_A+q_B)b(c^*)^2}{2}$$

$$cs = \frac{(q_A+q_B)r_i a_i^2}{8b}$$

The consumer surplus of all the advertisers is the sum of the surpluses of all advertisers

$$CS = \frac{1}{8b}\sum_{i=1}^{N^{cp}} r_i a_i^2$$

We can also write it in terms of n_p^{cp} and n_{np}^{cp}

$$CS = \frac{1}{8b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)$$

Total consumer surplus of internet users is defined by

$$CS^{nn} = CS_A^{nn} + CS_B^{nn}$$

Where CS_A^{nn} is the consumer surplus of users connected to AP A and CS_B^{nn} is the consumer surplus of users connected to AP B and

$$CS_B = 1 - CS_A$$

In case of NN internet users' consumer surplus is

$$CS_A^{nn} = \int_0^1 \int_0^{x_1^*} f_1(x) f_2(y)\ U_A^{nn}(x,y)\ dxdy$$

$$CS_B^{nn} = \int_0^1 \int_{x_1^*}^1 f_1(x) f_2(y)\ U_A^{nn}(x,y)\ dxdy$$

Total welfare is $TW = PS_{AP} + PS_{CP} + CS_{AD} + CS^{nn}$

$$TW = \frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{1}{4b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right) + \frac{1}{8b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right) + CS^{nn}$$

$$TW = \frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{3}{4b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right) + CS^{nn}(f_1(x), f_2(x), U^{nn}(x,y); v, \alpha, \beta, t, d)$$

Proof of Proposition 5

We defined each AP's profit under network diversity with both AP prioritizing traffic using the following expressions

$$\pi_A = (p_A + C_p^* T))\ q_A = (p_A + (\sum_{i=1}^{n_p^{cp}} \frac{a_i - T}{2b})T))\ q_A = (p_A + \sum_{i=1}^{n_p^{cp}} \frac{a_i T - T^2}{2b})\ q_A$$

Taking the derivative of π_A with respect to T and making this expression equal to zero (F.O.C) we calculate the equilibrium T for prioritized traffic

$$\frac{\partial \pi_A}{\partial T} = (\sum_{i=1}^{n_p^{cp}} \frac{a_i - 2T}{2b})\ q_A = 0 \rightarrow \frac{1}{2b}\left(\sum_{i=1}^{n_p^{cp}} a_i - 2\, n_p^{cp}\, T\right) = 0 \rightarrow T^* = \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2 n_p^{cp}}$$

Results are identical for AP B

If both APs offer prioritized traffic, the profit for each CP with prioritized traffic is defined by:

$$\pi_p^{cp} = r_i(q_A + q_B)\ (p^a(c)c\ - g(c)) = r_i(q_A + q_B)[(a_i - bc)c - Tc] = r_i(q_A + q_B)((a_i - T)\ c - bc^2)]$$

Taking the derivative of π_p^{cp} with respect to c and making this expression equal to zero (F.O.C.) we calculate the equilibrium quantity per CP for prioritized traffic

$$\frac{\partial \pi_p^{cp}}{\partial c} = r_i(q_A + q_B)\ ((a_i - T) - 2bc) = 0 \rightarrow c_p^* = \frac{a_i - T^*}{2b}$$

We defined each non-prioritized CP's profit with the following expressions

$$\pi_{np}^{cp} = p^a(c)c\ \ (q_A + q_B) r_i$$

Equilibrium quantity of content is defined by

$$\frac{\partial \pi_{np}^{cp}}{\partial c} = r_i(q_A + q_B)\ (a_i - 2bc) = 0 \rightarrow c_{np}^* = \frac{a_i}{2b}$$

Total prioritized content per internet user

$$C_p^* = \sum_{i=1}^{n_p^{cp}} \frac{a_i - T^*}{2b}$$

Total non-prioritized content per internet user

$$C_{np}^* = \sum_{i=1}^{n_{np}^{cp}} \frac{a_i}{2b}$$

Taking into account that $c_p^* = \frac{a_i - T^*}{2b}$ and $T^* = \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2 n_p^{cp}}$

$$c_p^* = \frac{a_i - T^*}{2b} = \frac{a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}}{2b} \quad \rightarrow \quad c_p^* = \frac{1}{2b}\left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]$$

Total prioritized content per internet user is defined by

$$C_p^* = \frac{1}{2b}\sum_{i=1}^{n_p^{cp}}\left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right] = \frac{1}{2b}\left[\sum_{i=1}^{n_p^{cp}} a_i - \frac{1}{2}\sum_{i=1}^{n_p^{cp}} a_i\right]$$

$$C_p^* = \sum_{i=1}^{n_p^{cp}} \frac{a_i}{4b}$$

Taking into account that $C_{np}^* = \sum_{i=1}^{n_{np}^{cp}} \frac{a_i}{2b}$ total content is

$$C^* = C_p^* + C_{np}^* = \frac{1}{2b}\left[\sum_{i=1}^{n_p^{cp}} \frac{a_i}{2} + \sum_{i=1}^{n_{np}^{cp}} a_i\right]$$

Proof of Proposition 6

Each CP profit in the networks that prioritize traffic is

$$\pi_p^{cp} = r_i(q_A + q_B)\,(p^a(c_p^*)c_p^* - g(c_p^*)) = r_i(q_A + q_B)\big((a_i - T)c_p^* - b(c_p^*)^2\big)]$$

$$\pi_p^{cp} = r_i(q_A + q_B)\left((a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}})\frac{1}{2b}\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right) - b\left(\frac{1}{2b}\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)\right)^2\right)$$

$$\pi_p^{cp} = r_i(q_A + q_B)\left(\frac{1}{2b}\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 - \frac{b}{4b^2}\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2\right)$$

$$\pi_p^{cp} = \frac{r_i(q_A + q_B)}{4b}\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2$$

Each CP profit in the networks that don't prioritize traffic where $c_{np}^* = \frac{a_i}{2b}$ is

$$\pi_{np}^{cp} = r_i p^a(c_{np}^*)c_{np}^* \;(q_A + q_B) = r_i(q_A + q_B)\big(a_i c_{np}^* - b(c_{np}^*)^2\big) = r_i(q_A + q_B)\left(\frac{a_i^2}{2b} - \frac{a_i^2}{4b}\right)$$

$$\pi_{np}^{cp} = \frac{(q_A + q_B)}{4b} r_i a_i^2$$

Total profit for all CPs is the producer surplus in the content provision market

$$PS^{cp} = PS_p^{cp} + PS_{np}^{cp} = \frac{(q_A + q_B)}{4b}\left[\sum_{i=1}^{n_p^{cp}} r_i\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right]$$

Considering that $q_A = q_B = \frac{1}{2}$

$$PS^{cp} = \frac{1}{4b}\left[\sum_{i=1}^{n_p^{cp}} r_i\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right]$$

Under network diversity if both AP prioritize traffic $\pi_A = (p_A + C_p^* T^*)\, q_A$ and $\pi_B = (p_B + C_p^* T^*)\, q_B$, $p_A^* = p_B^* = \frac{d}{f_1\left(\frac{1}{2}\right)}$, $C_p^* = \sum_{i=1}^{n_p^{cp}} \frac{a_i}{4b}$, $T^* = \frac{\sum_{i=1}^{n_p^{cp}} a_i}{n_p^{cp}}$ and $q_A = q_B = \frac{1}{2}$

$$PS_{AP} = \pi_A + \pi_B = p + C_p^* T^* = \frac{d}{f_1\left(\frac{1}{2}\right)} + \sum_{i=1}^{n_p^{cp}} \frac{a_i}{4b} \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}$$

$$PS_{AP} = \frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{1}{8bn_p^{cp}} \sum_{i=1}^{n_p^{cp}} a_i^2$$

We defined the consumer surplus of each advertiser per unit of content per user if the two networks offer prioritization for contents that pay for it as $cs_p = \frac{r_i(a_i - p^a(c_p^*;T^*))c_p^*\ (q_A+q_B)}{2}$ where $c^* = \frac{a_i - T^*}{2b}$ and $T^* = \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}$ therefore

$$cs_p = \frac{r_i(a_i - ((a_i - bc_p^*))c_p^*\ (q_A + q_B)}{2} = \frac{r_i(q_A + q_B)b(c_p^*)^2}{2} = \frac{r_i(q_A + q_B)(a_i - T^*)^2}{8b}$$

$$cs_p = \frac{(q_A + q_B)r_i(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}})^2}{8b}$$

$$CS_p = \frac{(q_A + q_B)}{8b}\left[\sum_{i=1}^{n_p^{cp}} r_i(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}})^2\right]$$

Considering that $q_A = q_B = \frac{1}{2}$

$$CS_p = \frac{1}{16b}\left[\sum_{i=1}^{n_p^{cp}} r_i(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}})^2\right]$$

We defined the consumer surplus of each advertiser per unit of content per user if the two networks offer prioritization for contents that do not pay for it as $cs_{np} = \frac{r_i(a_i - p^a(c_{np}^*))c_p^*\ (q_A+q_B)}{2}$ where $c^* = \frac{a_i}{2b}$ therefore

$$cs_{np} = \frac{r_i(a_i - (a_i - bc_{np}^*))c_{np}^*\ (q_A + q_B)}{2} = \frac{r_i(q_A + q_B)b(c_{np}^*)^2}{2}$$

$$cs_{np} = \frac{(q_A + q_B)r_i a_i^2}{8b}$$

The consumer surplus of all the advertisers is the sum of the surpluses

$$CS_{np} = \frac{(q_A + q_B)}{8b} \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2$$

Total consumer surplus is

$$CS = CS_p + CS_{np} = \frac{(q_A + q_B)}{8b} \left[\sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}} \right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2 \right]$$

Considering that $q_A = q_B = \frac{1}{2}$

$$CS = CS_p + CS_{np} = \frac{1}{16b} \left[\sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}} \right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2 \right]$$

Internet consumers consumer surplus in case of ND both offering prioritized traffic is

$$CS^{ndb} = CS_A^{ndb} + CS_B^{ndb}$$

Where

$$CS_A^{ndb} = \int_0^1 \int_0^{x_1^*} f_1(x) f_2(y)\, U_A^{ndb}(x,y)\, dxdy$$

$$CS_B^{ndb} = \int_0^1 \int_{x_1^*}^1 f_1(x) f_2(y)\, U_A^{ndb}(x,y)\, dxdy$$

Total welfare is $TW = PS_{AP} + PS_{CP} + CS_{AD} + CS^{ndb}$

$$TW = \frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{1}{4b} \left[\sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}} \right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2 \right] + \frac{1}{16b} \left[\sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}} \right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2 \right] + CS^{ndb}$$

$$TW = \frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{5}{16b} \left[\sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}} \right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2 \right] + CS^{ndb}$$

Proof of Proposition 7

AP B's profit is

$$\pi_B = p_B\ q_B\ = p_B\,[F_1(x_1) + \int_{x_1}^{x_2} (1 - F_2(y))\, f_1(x) dx\,] = p_B\,[F_1(x_1)] \int_{x_1}^{x_2}\ f_1(x) dx - \int_{x_1}^{x_2} (1 - F_2(y))\, f_1(x) dx$$

The marginal consumer located at $(x_{1,}0)$ and $(x_{2,}1)$ are defined by

$$\tilde{x_1} = \frac{(p_B - p_A) + (d - t\beta)}{2d} \qquad \widetilde{x_2} = \frac{(p_B - p_A) + (d + t\beta)}{2d}$$

Therefore

$$\frac{\partial x_1}{\partial p_A} = \frac{\partial x_2}{\partial p_A} = -\frac{1}{2d} \qquad \frac{\partial x_1}{\partial p_B} = \frac{\partial x_2}{\partial p_B} = \frac{1}{2d}$$

Taking the derivative of π_B with respect to p_B we obtain p_B^*. As we are looking for the pure strategy Bertrand-Nash equilibrium p_A^* will be equal to p_B^*

$$\frac{\partial\pi_B}{\partial p_B} = F_1(x_1) + \int_{x_1}^{x_2}(1 - F_2(y))\, f_1(x)dx \ + p_B \left[\ f_1(x_1)\frac{\partial x_1}{\partial p_B} + f_1(x_2)\frac{\partial x_2}{\partial p_B} - f_1(x_1)\frac{\partial x_1}{\partial p_B}\right.$$

$$\left. -(F_2(y = x_2)f_1(x_2)\frac{\partial x_2}{\partial p_B} - F_2(y = x_1)f_1(x_1)\frac{\partial x_1}{\partial p_B} + \int_{x_1}^{x_2} \frac{\partial F_2(y)}{\partial p_B} f_1(x)dx)\ \right] = 0$$

Taking into account that $F_2(y = x_2) = 1$ and $F_2(y = x_1) = 0$ and $F_1(x_1) + \int_{x_1}^{x_2}(1 - F_2(y))\, f_1(x)dx\ ) = \frac{1}{2}$

$$F_1(x_1) + \int_{x_1}^{x_2}(1 - F_2(y))\, f_1(x)dx \ - \ p_B \left[\ \int_{x_1}^{x_2} \frac{\partial F_2(y)}{\partial p_B} f_1(x)dx)\ \right] = 0$$

$$\frac{1}{2} - \frac{p_B}{2d}\left[\ \int_{x_1}^{x_2} f_2(y)f_1(x)dx)\ \right] = 0$$

Then we obtain

$$p_A^* = p_B^* = \frac{d}{\int_{x_1}^{x_2} f_2(y(x))f_1(x)dx}$$

Proof of Proposition 8

Being $\widetilde{x_1} = \frac{(d-t\beta)}{2d}$; $\widetilde{x_2} = \frac{(d+t\beta)}{2d}$ and $\tilde{y} = \frac{(d-t\beta)}{2t\beta} + \frac{d}{t\beta}x$ and $p_A^* = p_B^* = \frac{d}{\int_{x_1}^{x_2} f_2(x)f_1(x)dx}$ Let $A = \int_{x_1^*}^{x_2^*} f_2(y)f_1(x)dx$ then

$$\frac{\partial p^*}{\partial d} = \frac{\left(A - d\frac{\partial A}{d}\right)}{A^2} \quad and \quad \frac{\partial p^*}{\partial t} = \frac{\left(-d\frac{\partial A}{t}\right)}{A^2}$$

$$\frac{\partial A}{\partial d} = \left(f_2(y(x_2^*))f_1(x_2^*)\frac{\partial x_2^*}{\partial d} - f_2(y(x_1^*))f_1(x_1^*)\frac{\partial x_1^*}{\partial d}\right) + \int_{x_1^*}^{x_2^*} f_2'(y)\frac{\partial y}{\partial d} f_1(x)dx =$$

$$= \left(-f_2(y\left(\frac{d}{t\beta}\right))f_1(x_2^*)\frac{t\beta}{2d^2} - f_2\left(y(\frac{d}{t\beta} - 1)\right)f_1(x_1^*)\frac{t\beta}{2d^2}\right) + \int_{x_2^*}^{x_2^*} f_2'(y)\left(\frac{1}{2t\beta} - \frac{1}{t\beta}x\right)f_1(x)dx =$$

$$\frac{t\beta}{2d^2}\left(f_2\left(\frac{d}{t\beta}\right)f_1(x_2^*)\right) + \frac{1}{2t\beta}\int_{x_1^*}^{x_2^*} f_2'(y)(1 - 2x)f_1(x)dx$$

Since $y(x_2^*) = \frac{d}{t\beta}$; $y(x_1^*) = \frac{d-t\beta}{t\beta}$; $\frac{\partial x_2^*}{\partial d} = \frac{t\beta}{2d^2}$; $\frac{\partial x_1^*}{\partial d} = \frac{t\beta}{2d^2}$; $\frac{\partial y}{\partial d} = \frac{1}{2t\beta} - \frac{1}{t\beta}x$

$$\frac{\partial A}{\partial t} = \left(f_2(y(x_2^*))f_1(x_2^*)\frac{\partial x_2^*}{\partial t} - f_2(y(x_1^*))f_1(x_1^*)\frac{\partial x_1^*}{\partial t}\right) + \int_{x_1^*}^{x_2^*} f_2'(y)\frac{\partial y}{\partial t} f_1(x)dx =$$

$$= \left(f_2(y\left(\frac{d}{t\beta}\right))f_1(x_2^*)\frac{\beta}{2d} + f_2\left(y(\frac{d}{t\beta} - 1)\right)f_1(x_1^*)\frac{\beta}{2d}\right) + \int_{x_2^*}^{x_2^*} f_2'(y)\left(\frac{(t - 2t\beta)\beta - d}{2t^2\beta^2} - \frac{d}{t^2\beta}x\right)f_1(x)dx =$$

Since $y(x_2^*) = \frac{d}{t\beta}$; $y(x_1^*) = \frac{d-t\beta}{t\beta}$; $\frac{\partial x_2^*}{\partial t} = \frac{\beta}{2d}$; $\frac{\partial x_1^*}{\partial t} = -\frac{\beta}{2d}$; $\frac{\partial y}{\partial t} = \frac{(t-2t\beta)\beta - d}{2t^2\beta^2} - \frac{d}{t^2\beta}x$

Therefore

$$\frac{\partial p^*}{\partial d} = \frac{\int_{x_1^*}^{x_2^*} f_2(y) f_1(x) dx - \left(-f_2(y\left(\frac{d}{t\beta}\right)) f_1(x_2^*) \frac{t\beta}{2d^2} - f_2\left(y(\frac{d}{t\beta} - 1)\right) f_1(x_1^*) \frac{t\beta}{2d^2}\right) + \int_{x_2^*}^{x_2^*} f_2'(y) \left(\frac{1}{2t\beta} - \frac{1}{t\beta} x\right) f_1(x) dx}{(\int_{x_1^*}^{x_2^*} f_2(y) f_1(x) dx)^2}$$

$$\frac{\partial p^*}{\partial t} = \frac{\int_{x_1^*}^{x_2^*} f_2(y) f_1(x) dx - \left(f_2(y\left(\frac{d}{t\beta}\right)) f_1(x_2^*) \frac{\beta}{2d} + f_2\left(y(\frac{d}{t\beta} - 1)\right) f_1(x_1^*) \frac{\beta}{2d}\right) + \int_{x_2^*}^{x_2^*} f_2'(y) \left(\frac{(t - 2t\beta)\beta - d}{2t^2\beta^2} - \frac{d}{t^2\beta} x\right) f_1(x) dx}{(\int_{x_1^*}^{x_2^*} f_2(y) f_1(x) dx)^2}$$

Proof of Proposition 9

We defined each prioritized CP's profit with the following expressions

$\pi_A = p_A q_A$ does not depend on T

$$\pi_B = (p_B + C_p^* T)\, q_B = (p_B q_B + \frac{q_B T}{2b}(\sum_{i=1}^{n_p^{cp}} a_i - \frac{n_p^{cp} q_B T}{(q_A + q_B)}))$$

Taking the derivative of π_B with respect to T and making this expression equal to zero we can calculate the equilibrium T for prioritized traffic

$$\frac{\partial \pi_B}{\partial T} = \frac{q_B}{2b}\left(\sum_{i=1}^{n_p^{cp}} a_i - \frac{2n_p^{cp} q_B T}{(q_A + q_B)}\right) = 0 \rightarrow T^* = \frac{(q_A + q_B)\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}\, q_B}$$

We defined each prioritized CP's profit with the following expressions

$$\pi_p^{cp} = r_i(q_A + q_B)\, p^a(c)c - g(c)\, q_B = r_i(q_A + q_B)[(a_i - bc)c] - Tc\, q_B$$

Taking the derivative of π_p^{cp} with respect to c and making this expression equal to zero we can calculate the equilibrium quantity per CP for prioritized traffic

$$\frac{\partial \pi_p^{cp}}{\partial c} = r_i(q_A + q_B)[(a_i - 2bc)] - T\, q_B = 0 \rightarrow r_i(q_A + q_B)[(a_i - 2bc)] - T\, q_B = 0$$

$$a_i(q_A + q_B) - 2bc(q_A + q_B) - q_B T = 0$$

$$c_p^* = \frac{(q_A + q_B)a_i - q_B T^*}{2b(q_A + q_B)}$$

We defined each non-prioritized CP's profit with the following expressions

$$\pi_{np}^{cp} = r_i p^a(c)\, c\, (q_A + q_B)$$

Equilibrium quantity of content is defined by

$$\frac{\partial \pi_{np}^{cp}}{\partial c} = r_i(q_A + q_B)\,(a_i - 2bc) = 0 \rightarrow c^* = \frac{a_i}{2b}$$

Total prioritized content per internet user

$$C_p^* = \frac{1}{2b}(\sum_{i=1}^{n_p^{cp}} a_i - \frac{n_p^{cp}\ q_B T^*}{(q_A + q_B)})$$

Total non-prioritized content per internet user is

$$C_{np}^* = \sum_{i=1}^{n_{np}^{cp}} \frac{a_i}{2b}$$

Taking into account that $c_p^* = \frac{(q_A + q_B)a_i - q_B T^*}{2b(q_A + q_B)}$

$$c_p^* = \frac{(q_A\ +\ q_B)a_i - q_B T}{2b(q_A\ +\ q_B)} = \frac{(q_A\ +\ q_B)a_i - q_B\ \frac{(q_A\ +\ q_B)\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}\ q_B}}{2b(q_A\ +\ q_B)} = \frac{a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}}{2b}$$

$$c_p^* = \frac{1}{2b}\left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]$$

Total prioritized and non-prioritized content per internet user is

$$C_p^* = \frac{1}{2b}\sum_{i=1}^{n_p^{cp}}\left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right] = \frac{1}{2b}\left[\sum_{i=1}^{n_p^{cp}} a_i - \frac{1}{2}\sum_{i=1}^{n_p^{cp}} a_i\right] = \frac{1}{4b}\left[\sum_{i=1}^{n_p^{cp}} a_i\right] \qquad C_{np}^* = \sum_{i=1}^{n_{np}^{cp}} \frac{a_i}{2b}$$

Total traffic in all networks is

$$C^* = C_p^* + C_{np}^* = \frac{1}{2b}\left[\sum_{i=1}^{n_p^{cp}} \frac{a_i}{2} + \sum_{i=1}^{n_{np}^{cp}} a_i\right]$$

Proof of Proposition 10

Each CP profit in the networks that prioritize traffic is

$$\pi_p^{cp} = r_i(q_A\ +\ q_B)p^a(c_p^*)c_p^*\ - r_i g(c_p^*)\ q_B = r_i\left(q_A\big(a_i c_p^* - b(c_p^*)^2\big)\ +\ r_i q_B\big((a_i - T)c_p^* - b(c_p^*)^2\big)\right)$$

$$\pi_p^{cp} = r_i q_A\left(\frac{a_i}{2b}\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right) - \frac{1}{4b}\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2\right) + r_i q_B\left(\frac{1}{4b}\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2\right)$$

$$\pi_p^{cp} = \frac{r_i q_A}{4b}\left(2\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right) - \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2\right) + \frac{r_i q_B}{4b}\left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2$$

Each CP profit in the networks that don't prioritize traffic where $c_{np}^* = \frac{a_i}{2b}$ is

$$\pi_{np}^{cp} = r_i\, p^a(c_{np}^*)c_{np}^*\ (q_A + q_B) = r_i(q_A + q_B)\big(a_i c_{np}^* - b(c_{np}^*)^2\big) = r_i(q_A + q_B)\left(\frac{a_i{}^2}{2b} - \frac{a_i^2}{4b}\right)$$

$$\pi_{np}^{cp} = \frac{(q_A + q_B)}{4b} r_i a_i^2$$

Total profit of all CPs is the producer surplus in the content provision market

$$PS^{cp} = PS_p^{cp} + PS_{np}^{cp}$$

$$PS^{cp} = \frac{q_A}{4b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i{}^2 - \sum_{i=1}^{n_{np}^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2\right) + \frac{q_B}{4b}\sum_{i=1}^{n_{np}^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \frac{(q_A + q_B)}{4b}\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2$$

$$PS^{cp} = \frac{q_A}{4b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i{}^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i{}^2 - \sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2\right) + \frac{q_B}{4b}\left(\sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)$$

Considering that $q_A = q_B = \frac{1}{2}$

$$PS^{cp} = \frac{1}{4b}\left(\frac{1}{2}\sum_{i=1}^{n_p^{cp}} r_i a_i{}^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i{}^2\right)$$

Under network diversity if one AP prioritize traffic $\pi_A = p_A\, q_A$ and the other do not prioritize $\pi_B = (p_B + C_p^* T^*)\, q_B$, $p_A^* = p_B^* = \frac{d}{\int_{x_1}^{x_2} f_2(x) f_1(x) dx)}$, $C_p^* = \sum_{i=1}^{n_p^{cp}} \frac{a_i}{4b}$, $T^* = \frac{(q_A + q_B)\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}\, q_B}$ and $q_A = q_B = \frac{1}{2}$

$$\pi_A = p_A q_A$$

$$\pi_B = (p_B + C_p^* T)\, q_B$$

$$\pi_A = p_A q_A = \frac{d}{2\int_{x_1}^{x_2} f_2(x) f_1(x) dx)}$$

$$\pi_B = (p_B + C_p^* T)\, q_B = \frac{d}{2\int_{x_1}^{x_2} f_2(x) f_1(x) dx)} + \frac{1}{2}\sum_{i=1}^{n_p^{cp}} \frac{a_i}{4b}\left(\frac{(q_A + q_B)\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}\, q_B}\right)$$

$$PS_{AP} = \pi_A + \pi_B = \frac{d}{\int_{x_1}^{x_2} f_2(x) f_1(x) dx)} + \frac{1}{8bn_p^{cp}}\sum_{i=1}^{n_p^{cp}} a_i^2$$

We defined the consumer surplus of each advertiser per unit of content per user <u>if one networks offer prioritization for contents and the other don't</u> as $cs = \frac{(a_i - p^a(c_p^*;T^*))c_p^*\, q_B + (a_i - p^a(c_{np}^*))c_{np}^*\, q_A}{2}$ where $c_p^* = \frac{1}{2b}\left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]$ and $c_{np}^* = \frac{a_i}{2b}$ and $T^* = \frac{(q_A + q_B)\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}\, q_B}$ therefore

$$cs = \frac{r_i(a_i - p^a(c_p^*;T^*))c_p^*\, q_B + r_i(a_i - p^a(c_{np}^*))c_{np}^*\, q_A}{2} = \frac{r_i b}{2}\left[q_B (c_p^*)^2 + q_A (c_{np}^*)^2\right]$$

$$cs = \frac{br_i}{2}\left[q_B\left(\frac{1}{2b}\left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]\right)^2 + q_A\frac{r_i a_i^2}{4b^2}\right]$$

$$cs = \frac{r_i q_B \left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]^2 + 2q_A r_i a_i^2}{8b}$$

$$CS = \frac{1}{8b}\left(q_B \sum_{i=1}^{n_p^{cp}} r_i \left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]^2 + 2q_A \sum_{i=1}^{n_p^{cp}} r_i a_i^2\right)$$

We defined the consumer surplus of each advertiser per unit of content per user if one networks offer prioritization for contents and the other don't as $cs_{np} = \frac{r_i(a_i - p^a(c_{np}^*))c_p^*\ (q_A+q_B)}{2}$ where $c^* = \frac{a_i}{2b}$ therefore

$$cs_{np} = \frac{(q_A + q_B) r_i a_i^2}{8b}$$

and total consumer surplus of each advertiser per unit of content

$$CS_{np} = \frac{(q_A + q_B)}{8b}\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2$$

Total consumer surplus is

$$CS = CS_p + CS_{np} = \frac{1}{8b}\left[\left(q_B \sum_{i=1}^{n_p^{cp}} r_i \left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]^2 + (3q_A + q_B)\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)\right]$$

Considering that $q_A = q_B = \frac{1}{2}$

$$CS = \frac{1}{16b}\left[\left(\sum_{i=1}^{n_p^{cp}} r_i \left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]^2 + 4\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)\right]$$

Internet user's consumer surplus in case of ND where one AP offers prioritized traffic and the other non-prioritized traffic

$$CS^{ndo} = CS_A^{ndo} + CS_B^{ndo}$$

where

$$CS_A^{ndo} = \int_0^1 \int_0^{x_1^*} f_1(x) f_2(y)\, U_A^{ndo}(x,y)\, dxdy + \int_0^1 \int_{x_1^*}^{x_2^*} (1 - F_2(y(x))\, f_1(x)\, U_A^{ndo}(x,y)\, dxdy$$

$$CS_B^{ndo} = \int_0^1 \int_0^{x_1^*} F_2(y(x)) f_1(x)\, U_A^{ndo}(x,y)\, dxdy + \int_0^1 \int_{x_1^*}^{x_2^*} f_1(x) f_2(y)\, U_A^{ndo}(x,y)\, dxdy$$

Total welfare is $TW = PS_{AP} + PS_{CP} + CS_{AD} + CS^{ndo}$

$$TW = \frac{d}{\int_{x_1}^{x_2} f_2(x) f_1(x) dx} + \frac{1}{8bn_p^{cp}}\sum_{i=1}^{n_p^{cp}} a_i^2 + \frac{1}{4b}\left(\frac{1}{2}\sum_{i=1}^{n_p^{cp}} r_i {a_i}^2 + \sum_{i=1}^{n_{np}^{cp}} r_i {a_i}^2\right) + \frac{1}{16b}\left[\left(\sum_{i=1}^{n_p^{cp}} r_i \left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]^2 + 4\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)\right] + CS^{ndo}$$

Annex 1: Fundamental mathematical expressions for each scenario

Fundamental mathematical expressions for each scenario		
NN	**ND both prioritizing**	**ND one prioritizing**
$U_A = v\alpha - dx^2 - t\beta(y-1)^2 - p_A$	$U_A = v\alpha - dx^2 - t\beta y^2 - p_A$	$U_A = v\alpha - dx^2 - t\beta(y-1)^2 - p_A$
$U_B = v\alpha - d(1-x)^2 - t\beta(1-y)^2 - p_B$	$U_B = v\alpha - d(1-x)^2 - t\beta y^2 - p_B$	$U_B = v\alpha - d(1-x)^2 - t\beta y^2 - p_B$
$x_1 = \frac{1}{2} \quad x_2 = \frac{1}{2}$	$x_1 = \frac{1}{2} \quad x_2 = \frac{1}{2}$	$\widetilde{x_1} = \frac{(d-t\beta)}{2d} \quad \widetilde{x_2} = \frac{(d+t\beta)}{2d}$
$p \equiv p_A^* = p_B^* = \frac{d}{f_1\left(\frac{1}{2}\right)}$	$p \equiv p_A^* = p_B^* = \frac{d}{f_1\left(\frac{1}{2}\right)}$	$p^* \equiv p_A^* = p_B^* = \frac{d}{\int_{x_1}^{x_2} f_2(y) f_1(x)dx}$
$p^a(c^*) = a_i - bc^*$	$p^a(c^*) = a_i - bc^*$	$p^a(c^*) = a_i - bc^*$
	$T^* = \frac{\sum_{i=1}^{n_p^{cp}} a_i}{n_p^{cp}}$	$T^* = \frac{(q_A + q_B)\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}\, q_B}$
$c^* = \frac{a_i}{2b}$	$c_p^* = \frac{1}{2b}\left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]$	$c_p^* = \frac{1}{2b}\left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]$
	$c_{np}^* = \frac{a_i}{2b}$	$c_{np}^* = \frac{a_i}{2b}$
	$C_p^* = \sum_{i=1}^{n_{np}^{cp}} \frac{a_i}{4b}$	$C_p^* = \sum_{i=1}^{n_{np}^{cp}} \frac{a_i}{4b}$
$C^* = \sum_{i=1}^{N^{cp}} \frac{a_i}{2b}$	$C^* = \frac{1}{2b}\left[\sum_{i=1}^{n_p^{cp}} \frac{a_i}{2} + \sum_{i=1}^{n_{np}^{cp}} a_i\right]$	$C^* = \frac{1}{2b}\left[\sum_{i=1}^{n_p^{cp}} \frac{a_i}{2} + \sum_{i=1}^{n_{np}^{cp}} a_i\right]$

Results for producer surplus of access and content providers and consumer surplus of advertisers and users		
	$\boldsymbol{PS_{AP}}$	
NN	**ND both prioritizing**	**ND one prioritizing**
$\frac{d}{f_1\left(\frac{1}{2}\right)}$	$\frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{1}{8bn_p^{cp}} \sum_{i=1}^{n_{np}^{cp}} a_i^2$	$\frac{d}{\int_{x_1}^{x_2} f_2(y(x)) f_1(x)dx} + \frac{1}{8bn_p^{cp}} \sum_{i=1}^{n_{np}^{cp}} a_i^2$
	$\boldsymbol{PS_{CP}}$	
NN	**ND both prioritizing**	**ND one prioritizing**
$\frac{1}{4b}\left[\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right]$	$\frac{1}{4b}\left[\sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right]$	$\frac{1}{4b}\left[\frac{1}{2}\sum_{i=1}^{n_p^{cp}} r_i {a_i}^2 + \sum_{i=1}^{n_{np}^{cp}} r_i {a_i}^2\right]$
	$\boldsymbol{CS_A}$	
NN	**ND both prioritizing**	**ND one prioritizing**
$\frac{1}{16b}\left[\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right]$	$\frac{1}{16b}\left[\sum_{i=1}^{n_p^{cp}} r_i (a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}})^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right]$	$\frac{1}{16b}\left[\left(\sum_{i=1}^{n_p^{cp}} r_i \left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]^2 + 4\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)\right]$
	$\boldsymbol{CS_A}$	
NN	**ND both prioritizing**	**ND one prioritizing**
$CS_A^{nn} = \int_0^1 \int_0^{x_1^*} f_1(x) f_2(y)\, U_A^{nn}(x,y)\, dxdy$ $CS_B^{nn} = \int_0^1 \int_{x_1^*}^{x_2^*} f_1(x) f_2(y)\, U_A^{nn}(x,y)\, dxdy$	$CS_A^{ndb} = \int_0^1 \int_0^{x_1^*} f_1(x) f_2(y)\, U_A^{ndb}(x,y)\, dxdy$ $CS_B^{ndb} = \int_0^1 \int_{x_1^*}^{x_2^*} f_1(x) f_2(y)\, U_A^{ndb}(x,y)\, dxdy$	$CS_A^{ndo} = \int_0^1 \int_0^{x_1^*} f_1(x) f_2(y)\, U_A^{ndo}(x,y)\, dxdy$ $+ \int_0^1 \int_0^{x_1^*} (1 - F_2(y(x))\, f_1(x)\, U_A^{ndo}(x,y)\, dxdy$ $CS_B^{ndo} = \int_0^1 \int_0^{x_1^*} F_2(y(x)) f_1(x)\, U_A^{ndo}(x,y)\, dxdy +$ $\int_0^1 \int_{x_1^*}^{x_2^*} f_1(x) f_2(y)\, U_A^{ndo}(x,y)\, dxdy$

Difference in total welfare between network neutrality regulation and network diversity, both APs offering prioritized traffic
$\nabla TW^{NN/NDB} = \frac{3}{4b}\sum_{i=1}^{n_p^{cp}} r_i a_i^2 - \sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \frac{7}{16b}\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2 + CS^{nn} - CS^{ndo}$
Difference in total welfare between network neutrality regulation and network diversity, one AP offering prioritized traffic and the other offering network neutrality
$\nabla TW^{NN/NDB} = \frac{d}{f_1\left(\frac{1}{2}\right)} - \frac{d}{\int_{x_1}^{x_2} f_2(x) f_1(x) dx)} + \frac{1}{2b}\left(\frac{1}{2}\sum_{i=1}^{n_p^{cp}} r_i {a_i}^2 + \sum_{i=1}^{n_{np}^{cp}} r_i {a_i}^2\right) - \frac{1}{8bn_p^{cp}}\sum_{i=1}^{n_p^{cp}} a_i^2 - \frac{1}{16b}\left[\left(\sum_{i=1}^{n_p^{cp}} r_i \left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]^2 + 4\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)\right] + CS^{nn} - CS^{ndo}$

The difference in total welfare between network neutrality regulation and network diversity, both APs offering prioritized traffic is:

$$\nabla TW^{NN/NDB} = TW^{NN} - TW^{NDB} = \frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{3}{4b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right) + CS^{nn} - \left[\frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{5}{16b}\left[\sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right] + CS^{ndb}\right] = \frac{3}{4b}\sum_{i=1}^{n_p^{cp}} r_i a_i^2 - \sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \frac{7}{16b}\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2 + CS^{nn} - CS^{ndo}$$

$$\nabla TW^{NN/NDB} = \frac{3}{4b}\sum_{i=1}^{n_p^{cp}} r_i a_i^2 - \sum_{i=1}^{n_p^{cp}} r_i \left(a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right)^2 + \frac{7}{16b}\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2 + CS^{nn} - CS^{ndo}$$

The difference in total welfare between network neutrality regulation and network diversity, one APs offering prioritized traffic and the other offering network neutrality is:

$$\nabla TW^{NN/NDB} = TW^{NN} - TW^{NDO} = \frac{d}{f_1\left(\frac{1}{2}\right)} + \frac{3}{4b}\left(\sum_{i=1}^{n_p^{cp}} r_i a_i^2 + \sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right) + CS^{nn} - \left[\frac{d}{\int_{x_1}^{x_2} f_2(x) f_1(x) dx)} + \frac{1}{8bn_p^{cp}}\sum_{i=1}^{n_p^{cp}} a_i^2 + \frac{1}{4b}\left(\frac{1}{2}\sum_{i=1}^{n_p^{cp}} r_i {a_i}^2 + \sum_{i=1}^{n_{np}^{cp}} r_i {a_i}^2\right) + \frac{1}{16b}\left[\left(\sum_{i=1}^{n_p^{cp}} r_i \left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]^2 + 4\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)\right] + CS^{ndo}\right]$$

$$\nabla TW^{NN/NDB} = \frac{d}{f_1\left(\frac{1}{2}\right)} - \frac{d}{\int_{x_1}^{x_2} f_2(x) f_1(x) dx)} + \frac{1}{2b}\left(\frac{1}{2}\sum_{i=1}^{n_p^{cp}} r_i {a_i}^2 + \sum_{i=1}^{n_{np}^{cp}} r_i {a_i}^2\right) - \frac{1}{8bn_p^{cp}}\sum_{i=1}^{n_p^{cp}} a_i^2 - \frac{1}{16b}\left[\left(\sum_{i=1}^{n_p^{cp}} r_i \left[a_i - \frac{\sum_{i=1}^{n_p^{cp}} a_i}{2n_p^{cp}}\right]^2 + 4\sum_{i=1}^{n_{np}^{cp}} r_i a_i^2\right)\right] + CS^{nn} - CS^{ndo}$$

Annex 2: Simulation values and ranges for parameters value

Number of content providers offering prioritized content $n_p^{cp} = 50$

Number of content providers offering non-prioritized content $n_{np}^{cp} = 50$

Utility of internet consumption $v = 1$

Marginal value that a consumer places on internet access waiting time per unit $t = 0.12$

Marginal value that a consumer places on internet access $d = 0.06$

Ranges for the slope of advertisement inverse demand curve parameter $b \in [0.4, 1]$

Average value for $a_i = 0.325$

Average value for $r_i = 0.01$

Concentrated sigma: $\sigma = 0.25$

Disperse sigma: $\sigma = 0.75$

Number of simulations: 10

Normal distribution parameters $\mu = 0.5$; $\sigma = 0.25$

References

Armstrong, M. (2006). Competition in two-sided markets. *RAND Journal of Economics*, 37, 668–691.

Bourreau, M., Kourandi, F. and Valletti, T. (2015). Net Neutrality with Competing Internet Platforms. *The Journal Industrial Economics*, 63: 30–73. doi:10.1111/joie.12068

Caplin, A., Nalebuff, B. (1991). Aggregation and imperfect competition: On the existence of equilibrium. *Econometrica* 59, 25-59

Chen, M. Keith and Nalebuff, Barry J., One-Way Essential Complements (2006). Cowles Foundation Discussion Paper No. 1588; Yale Economic Applications and Policy Discussion Paper No. 22.

Cheng, H. K., S. Bandyopadhyay, H. Guo. (2011). The debate on net neutrality: A policy perspective. *Information Systems Research* 22(1): 60–82.

Choi, J. P., Kim,B. (2010). Net neutrality and investment incentives. *RAND Journal of Economics* 41(3): 446–471.

Choi, J.P., Jeon, D., Kim, B. (2013). Asymmetric Neutrality Regulation and Innovation at the Edges: Fixed vs Mobile Networks. NET Institute Working Paper No. 13-24.

Clarke, R. (2009). Costs of neutral/unmanaged IP networks. *Review of network economics*, Vol 8, Issue 1.

D'Aspremont, C., Gabszewicz, J.,Thisse, J. (1979). Hotellings Stability in Competition. *Econometrica* 47(5): 1145-1150

Economides, N. (1986). Minimal and Maximal Product Differentiation in Hotelling Duopoly. *Economics Letters* 21(1): 67-71.

Economides, N., Hermalin, B. (2012). The economics of network neutrality. *RAND Journal of Economics*. Vol 43 N4 pp 602-629.

Economides, N. Hermalin, B. E. (2015). The strategic use of download limits by a monopoly platform. *RAND Journal of Economics*, 46: 297–327. doi:10.1111/1756-2171.12087

Economides, N., Tag, J. (2012). Net neutrality on the internet: A two-sided market analysis. Inf. Econ. Policy 24(2): 91–172.

Faulhaber, G. (2011). The economics of network neutrality: a review. *Communications & Convergence Review*, 2011, Vol. 3, No.1, 53-64

Ford, G. Spiwak, E. (2010). Burden of Network Neutrality mandates on rural broadband deployment. *The journal of applied economy.* 238.

García-Zaballos, A., González-Herranz, F. (2013). From cybersecurity to cybercrime: A framework for analysis and implementation. Inter-American Development Bank. Technical Note Nº. IDB-TN-588

Hermalin, B. E., Katz, M.L. (2007). The economics of product-line restrictions with an application to the network neutrality debate. *Information Economics and Policy* 19: 215–248.

Hemphill, C. (2008). Network Neutrality and the false promise of zero price regulation. *Yale Journal on Regulation*. Vol 25:2, 2008.

Hotelling, H. (1929). Stability in competition. *The Economic Journal*, 39(153): 41–57

Huergo, E. and Muñoz, I.J. (2004). Condicionantes de la competencia en los servicios de telecomunicaciones: Evidencia internacional. *Revista de economía industrial,* Num. 361. 59-68

Irmen, A. and Thisse, J. (1998). Competition in Multi-characteristics Spaces: Hotelling Was Almost Right. *Journal of Economic Theory* 78: 76-102

Lee, D., Hwang, J. (2011). Network neutrality and difference in efficiency among Internet application service providers: A meta- frontier analysis. *Telecommunications Policy*, 35 (2011) 764-772.

Liu, Q. Shuai, J. (2015). On multi-Dimensional Product Differentiation with non-uniform distribution. Working paper. Available at SSRN: http://ssrn.com/abstract=1975191 or http://dx.doi.org/10.2139/ssrn.1975191.

Mialon, S. H. Banerjee, S. (2013). Regulation of Internet Access (August 1, 2013). Available at SSRN: http://ssrn.com/abstract=1965375 or http://dx.doi.org/10.2139/ssrn.1965375

Njoroge, P., Ozdaglar, A., Stier-Moses, N.E., Weintraub, G.Y. (2013). Investment in Two-Sided Markets and the Net Neutrality Debate. *Review of Network Economics,* 2013; 12(4): 355–402

Nocke, V., Peitz, M., Stahl, K., (2007). Platform ownership. Journal of the European Economic Association 5, 1130–1160.

Peitz, M. and Schuett, F. (2015). Net Neutrality and Inflation of Traffic (March 3, 2015). TILEC Discussion Paper No. 2015-006. Available at SSRN http://ssrn.com/abstract=2573466 or http://dx.doi.org/10.2139/ssrn.2573466

Porter, M. (1991). Towards a dynamic theory of strategy. *Strategic management journal*, Vol.12 95-117

Rochet,J., Tirole, J.(2003). Platform competition in two-sided markets. *Journal of the European Economic Association* 1, 990–1029.

Rochet, J., Tirole, J. (2006). Two-sided markets: a progress report. *RAND Journal of Economics,* 37, 645–667.

Shin, D.H. (2014). A comparative analysis of net neutrality: Insights gained by juxtaposing the US and Korea. *Telecommunications Policy*, 38 (2014) 1117–1133

Tabuchi, T. (1994). Two-stage two-dimensional spatial competition between two firms. *Regional Science and Urban Economics,* 24: 207-227.

Van Schewick, B. (2006), Towards an Economic Framework for Network Neutrality Regulation. Journal on *Telecommunications & High Technology Law*, 5, pp. 329–392.

Weyl, G. (2010). A price theory of multi-sided platforms. *American Economic Review*, 100, 1642–1672

Wright, D. (2013). Broadband Policy, consumer welfare. The case for antitrust approach to net neutrality issues. Information economy's projects conference on US broadband markets.

Wu, T., Yoo, C. (2006). Keeping Internet neutral? Tim Wu and Christofer Yoo Debate. *Federal Communications law journal*. Vol 59. Number 3.

Notes

[1] Traditionally know as ISP (Internet Service Provider)

[2] https://www.sandvine.com/downloads/general/global-internet-phenomena/2014/1h-2014-globalinternet-phenomena-report.pdf

[3] A comparative analysis of the evolution of the market size of each segment of the internet value chain from 2010 to 2016 can be found here:
https://www.atkearney.com/documents/10192/178350/internet-value-chain-economics.pdf/bd910b2c-bdae-4d6f-8903-f5edad6784eb
http://www.gsma.com/publicpolicy/wp-content/uploads/2016/05/GSMA_The-internet-Value-Chain_WEB.pdf

[4] See: http://ec.europa.eu/futurium/en/content/strengthening-consumer-protection-and-ensuring-consistent-rules

[5] See Articles 76 and 77, Protecting and Promoting the Open Internet, Report and Order on Remand, Declaratory Ruling, and Order, FCC 15-24 (Mar. 12, 2015). https://www.fcc.gov/document/fcc-releases-open-internet-order

[6] http://www.justice.gov/atr/us-v-microsoft-courts-findings-fact#vf

[7] AT&T sued Vonage on patent infringement. Vonage reached an agreement with AT&T. More information on this case can be found here.
http://www.sec.gov/Archives/edgar/data/1272830/000119312508059036/dex1039.htm . A similar case that also shows a refusal to deal case can be found at
https://transition.fcc.gov/ogc/documents/opinions/2004/02-682-011304.pdf

[8] See, https://ecfsapi.fcc.gov/file/60001031582.pdf

[9] Multi-homing refers to the availability of multiple internet APs in the area a consumer or a content provider is located.

[10] https://apps.fcc.gov/edocs_public/attachmatch/FCC-15-24A1.pdf

[11] Competition regulation might also change market structure, especially in low density population areas, for example mandating to share the access network with an entrant or forcing the existence of mobile network virtual operators.

[12] For example, technology improvements in data transmission and compression technologies reduce network congestion and the perception of increased quality of prioritized traffic. Also, the scarcity of spectrum restricts the number of mobile broadband providers restricting ND to only one network.

[13] A super- index of the type NN, NB or NO in any of the consumer and producer surplus or total welfare expression identifies the scenario being analyzed.

[14] It should be noted that the increased utility of reduced delay is captured in the model by the position of the AP in the square. Some positions have higher utility for consumers that value the most prioritization.

[15] For example, an advertiser will pay more to a content provider the higher the number of internet users consuming the content. Similarly, an advertiser will pay more to a content provider the higher the quantity of content being distributed to the internet user. For example, video streaming has higher quality the higher the number of data packages being transmitted keeping the rest of variables, such as compression technology, constant.

[16] For example Google, the market leader in internet advertising and marketing, has created the service AdSense for advertisers. This service offer advertisement space to publicist in internet sites. AdSense works crawling web pages for content and delivering ads based on that content. The more rich and relevant is the content the more possibilities have the site to receive higher income from advertisers. The price of and ad space is calculated using a cost per click bid and a score metric evaluating the quality of the ad. The quality score measures how useful it is the ad to the people who see it in this site and it is based on several factors, including the predicted click-through rate (performance of the ad on a site), the relevance of the advertisers and the set of keywords of the site.

[17] See:
http://www.digitalstrategyconsulting.com/intelligence/2014/06/ad_revenue_per_user_google_facebook_twitter.php

http://www.businessinsider.com/facebook-average-revenue-per-user-2015-11

CONCLUSIONS

The digital revolution is deeply transforming the economy and society. It has created enormous opportunities and also new risks. Growth, enhanced productivity, better democracy and social inclusion are some of the opportunities but at the same time these transformations are challenging human and consumer rights, social policies and economic efficiency.

Recent literature on digital economy issues is focused on three relevant fields. The first is the optimal supply of digital networks with externalities and high entry barriers. Barriers stem from the economics of deployment of digital networks, especially some network segments, and from the restricted access to the radio-electric spectrum. Secondly, the creation of market power linked to the concentrated sale of digital goods and services in a few vendors. Market power is generated from the demand side due to network externalities and reduced costs of digital transactions. These two factors are leading to the creation of monopsony power that may affect not only the new digital but also traditional markets. Finally, the transactions and interactions among network providers, digital service providers, consumers and vendors raise difficult questions about the optimal prices to remunerate efficiently all involved parties.

This thesis proposes regulations and strategies for the economic efficiency and optimal provision of digital networks, goods and services contributing to some of the topics of focus in recent literature on digital economy. One problem is the existence of entry barriers to companies providing internet access. We contribute to the literature by analyzing efficient allocation of spectrum, one of its causes, in chapters 1 and 2. Another problem is the optimal prices to remunerate network providers, digital service providers, consumers and vendors. We contribute by examining the impact of price discrimination of internet service providers under duopolistic competition and multi-dimensional product differentiation in retail broadband access in chapter 3.

Chapter 1 offers some proposals for the evaluation of the efficient allocation of spectrum to radio communication services. New approaches to spectrum management have resulted in a more efficient production of services. These approaches enabled better quality of service, the reduction of the amount of spectrum required, increased coverage and reduced prices. However, the role of the public sector is still essential in spectrum allocation. Technology advancement and improvements in the licensing methods must be complemented with processes of reallocation led by the public sector enabling higher allocative efficiency of resources. This chapter provides a methodology to measure the net benefit of the reallocation of a spectrum band intended to guide regulators and policy makers. After the reallocation, the band will be used by the incumbent and new entrant services. We have identified the following facts.

- The calculation of benefits and of the opportunity costs of spectrum usage should include the external value associated to the provision of services in cases where externalities can only be obtained using the studied frequency band.

- Demand for spectrum is driven by the quantity and quality of services consumed. For some services the population density is also a driver of spectrum demand.

- The cost of deploying networks to use the spectrum is a key element to determine the private value associated to the use of a frequency band. Deployment costs may be approximated by using production functions. The production function should be composed by several input arguments one of which is the spectrum. There are other inputs such as the number of stations, and the use of wired solutions that may be substitutes of a frequency band. If the service is using other frequencies all of them should be included in the analysis.

- Utility of spectrum usage is variable. It depends on the portion of spectrum being considered. Some portions are in great demand and have higher prices. Different portions of spectrum are not perfect substitutes one another, consequently, frequencies have variable substitution margins. The allocation of frequencies with similar features to services produces decreasing marginal returns. The key goal of spectrum management is to allocate the spectrum so that the marginal benefit obtained through the allocation of spectrum to an entrant equals the marginal opportunity cost of the incumbent relinquishing a portion of spectrum.

- Reallocation of spectrum produces transition costs that should be included in the cost benefit analysis. The cost of transition may include the adaptation of transmitting and/or receiving devices to change the frequencies used. Operating expenses stemming from the provision of services in the new and old frequencies to avoid service interruption during the transition period, costs of minimizing harmful interferences during the deployment of the entrant's network and costs associated to coverage losses. Sometimes the reallocation may also reduce costs if the provision of the incumbent service is streamlined.

- If a band is planned to be allocated in a harmonized way across a large geographical area with high expected number of customers, economies of scale can be obtained in the manufacture of the equipment required to use the frequency band. There may also be economies of scale in service provision.

Chapter 2 describes the optimal path and speed of spectrum management reform under associated uncertain costs and benefits. The question in spectrum management is no longer if reform is necessary to give market more participation but the path and speed of reform with current and expected technology advance. In this chapter, we analyze reform types categorizing them in terms of the type of efficiency gain that the reform is intended to produce and the associated uncertainty stemming from the intrinsic physical features of spectrum and the insufficient internalization of technology advancements in regulation. The chapter offers a model to ascertain the optimal speed of reform and shed some light on the policy options to reduce uncertainty in the associated costs and benefits of reform. We have determined when to choose a gradual or a big bang reform depending on technology and whether to wait or not for a new technological advancement. Both results depend on the expected outcome of the reform, the probability of realization of the states of nature and the value of reversal to the previous state of nature.

Transactions enabling allocative and dynamic efficiency are not yet possible to be produced solely by market forces. New approaches such as the design of incentive auctions give a broader role to market participants in spectrum allocation but the active contribution of the regulatory authority is still necessary to design the direct and reverse auction and the repackaging process. A different instrument, new technologies enabling to share the spectrum previously used by an incumbent service, also improve allocative efficiency making market agents more involved in spectrum allocation. However, sharing techniques are usually limited to either short range devices, in the case of underlay spectrum techniques, or by the availability of unused spectrum in certain areas or time slots, the case of dynamic spectrum sharing. Spectrum sharing techniques are not means to enable the change of use of a band solely by market forces.

Chapter 3 presents a model of price discrimination of over the top providers under duopolistic competition and multi-dimensional product differentiation in retail broadband access. Network neutrality regulation prevents price discrimination from access providers to content providers and product differentiation in terms of connection quality in the retail broadband access market. This chapter analyzes the economic implications of price discrimination under duopolistic competition and multi-dimensional product differentiation in retail internet access using a sequential-moves game theoretic model. Under this framework, we discuss the impact of product differentiation and price discrimination on social welfare, and offer systematic simulations using feasible ranges for parameters value to help discern the impact of departing from network neutrality regulation on social welfare.

Abandoning network neutrality regulation results in a reduction of the quantity of internet content produced. The new regulatory approach, network diversity, produces a significant reduction of prioritized content while the quantity of content non-susceptible of being prioritized remains unchanged. The reduction in the quantity of content is equal in the two network diversity scenarios considered - both and only one access provider prioritizing traffic - thus, even when only one access provider offers prioritized content the reduction is materialized on equal footing.

Under network diversity, access providers enjoy higher income from the tax charged to CPs for prioritization but content providers offering prioritized content, and advertisers suffer an income reduction. Income reduction depends on the sum of prices paid by advertisers to content providers and also on consumer preferences when only one access provider offer prioritization. Firms supplying non-prioritized content enjoy the same surplus. Abandoning network neutrality has two different types of effects on consumer's utility. On one hand, it reduces utility through a decrease in the value of the ratio of content actually available to content available with perfect competition. On the other hand, the disutility caused by the delay of data packages is reduced via the ratio of prioritized content to content not susceptible of prioritization. The simulation is necessary to measure net impact. A change in the regulatory regime would lead to different internet access market structures because there is a rotation of the indifferent consumer´s line via changes in the ratio of prioritized content to content not susceptible of prioritization. Even when the marginal disutility that consumers place on product options and waiting time remains constant, the change in the ratio of prioritized content to content not susceptible of prioritization would rotate the indifferent consumer´s line.

Results are expressed with respect to the maximum utility that a consumer could obtain with perfect competition in the provision of internet content, which is set to 1. However, consumers don't enjoy the maximum utility but a portion of it, because there exists reduced competition and/or a tax on content production.

Simulations show that network neutrality regulation is welfare superior to network diversity under the model assumptions with the values and value ranges given to parameters. Departing from network neutrality regulation leads to an abrupt decrease of the consumer surplus of internet users, from 0.4 to between 0.01 and 0.26. This effect weights much more in total welfare than the increase in APs surplus coming from the tax charged to content providers.

The increase in the producer surplus of access providers, from 0.06 to a maximum of 0.34 is not enough to compensate the reduction of the consumer surplus of internet users. Among the two network diversity scenarios, results are very similar with slightly higher total welfare in the case where only one content provider offers prioritized traffic. Disperse range of parameter values for content's relevance and the sum of prices paid by all advertisers tend to produce higher total welfare.

We can set a monetary value for the maximum utility parameter v to make simulation results more intuitive. For example, if we set the value to 150$ per month, the simulation shows that total welfare would range from 57$ to 102$, consumer's utility from 1$ to 60$. The average ARPU for access providers varies from 9$ to 51$, for CPs from 8$ to 20$, and for advertisers from 2$ to 6.8$. Although the values for content's relevance and the sum of prices paid by all advertisers have been generated randomly using Montecarlo simulation modeling, they have a significant relation with real observed ARPU values.

Although additional extensions can be found in each chapter, we point out some extensions to the studies carried out in this thesis. For example, the first block could be complemented with the empirical analysis of the value of spectrum re-allocation. For some spectrum bands, the price paid in an auction is a good proxy for the market value of the portion of spectrum being auctioned. In the second block, another field left for future research is the analysis of the potential impact of a change on the internet access pricing regulation on the incentives to invest and innovate. There might be an impact in both, the provision of digital assets and the deployment of networks.

REGULATION

Table of Regulation.17: Regulation of the digital transformation in the European Union

Economic Policy Area	Topic	Legislative Development	Scope
Competition Spectrum allocation	General regulatory Framework for ECN and associated services	Directive 2002/21/EC Directive 2009/140/EC Regulation 544/2009	Defining procedures for: - Spectrum allocation secondary market, allocation of numbers - Standardization and interoperability -Security of networks - Ex-ante ECN regulation
Competition	Access and interconnection to ECN and associated facilities	Directive 2002/19/EC Directive 2009/140/EC	Defining obligations of ECN providers with significant market power
Authorization of ECN and access to rights of way Spectrum usage rights	Authorization of ECN networks and services	Directive 2002/20/EC Directive 2009/140/EC	Defining authorization rules and conditions for : -ECN deployment -Rights of use of spectrum/Rights of use of numbers
Digital social policy-externalities Network neutrality	Universal Service Consumer rights Network neutrality	Directive 2002/22/EC Directive 2009/136/EC Regulation 2015/2120	Rules and conditions for: -Universal service obligations -Consumer's rights protection -Equal and non-discriminatory treatment of internet traffic
Digital fundamental rights	Privacy and confidentiality of personal data	Directive 2002/58/EC Directive 2006/24/EC Directive 2009/136/EC	Protection of fundamental rights and freedoms in the digital world. Privacy and confidentiality
Competition	Establishment of the regulatory body for ECN, BEREC.	Regulation 1211/2009 (New regulation forthcoming)	Responsibilities of BEREC -Disseminate best practices -Regulatory assistance -Issue reports
Competition	Regulation of prices of the access to ECN when abroad	Regulation 531/2012 Regulation 2015/2120	Ensuring users do not pay excessive prices using ECN when travelling around Europe
Spectrum allocation	Establishment of the Radio Spectrum Policy Group	Decision 2002/622/EC Decision 2009/978/EU	Responsibilities of RSPG: -Assistance to the Commission on spectrum policy - Preparation of multiannual radio spectrum policy programs -Spectrum use harmonized conditions
Reallocation of spectrum-technology change	Spectrum allocation and technology change	Directive 2009/114/EC Council Directive 87/372/EEC Decision 2016/687 Decision 2016/339 Decision 2016/750	-Technological neutrality of spectrum use -Harmonized use of spectrum in Europe
Digital social policy-externalities	Public aid for connectivity in public spaces	Wi-Fi for all regulation forthcoming	Public Aid. Free Wi-Fi in public spaces
Competition Spectrum allocation Authorizations and usage rights	Unified regulatory Framework for ECN	European code for electronic communications forthcoming	Unification of framework, access, authorization and universal service directives

Source: Authors

Table of Regulation.18: Regulation of the digital transformation in Spain

Economic Policy Area	Topic	Legislative Development	Scope
Competition Property rights definition Consumer's rights Digital constitutional rights	General regulatory framework for Telecommunications	Ley 9/2014 General de Telecomunicaciones	General principles for: - Competition ECN - Universal service - ECN access to rights of way - Consumer's data privacy - Consumer rights -Standardization of devices -Spectrum property rights
Digital transactions Network neutrality	Consumer rights, digital services and e-commerce	Ley 34/2002 Orden PRE/361/2002	Conditions for digital service provision: -Consumer's data storage -Digital contracts -Digital advertising -Respect for fundamental rights (intellectual property, health protection, non-discrimination)
Contract obligations for ECN services	Consumer rights ECN	Real Decreto 899/2009	Access to ECN services conditions - Quality of service -Contract obligations
Digital social policy-Externalities	Universal service Obligations	Real Decreto 424/2005	Conditions -Access to rights of way -Universal service obligations -Consumer's rights
Competition	Establishment of the regulatory body for competition, CNMC	Ley 3/2013 Real Decreto 657/2013	CNMC responsibilities on digital markets: -Analysis and definition of markets with ex ante regulation -Identification of significant market power (SMP) -Conditions for companies with SMP -Conflict resolution
Competition among ECN providers	Regulation of ECN markets	Real Decreto 2296/2004	Conditions -Ex-ante regulation of ECN -Access and interconnection to ECN - Plans for allocation of numbers
Property rights. Secondary market Spectrum allocation plans	Spectrum use	Real Decreto 844/1989 Real Decreto 123/2017 Orden IET/1311/2013	Regulation of spectrum -Property rights -Secondary market -Spectrum allocation plans
Spectrum allocation table	Spectrum Allocation Table in Spain	Orden IET/787/2013	Table of frequency allocations, international and differential uses in Spain
Reallocation of spectrum-technology change	Spectrum allocation and technology change	Real Decreto 1773/1994	-Technological neutrality in spectrum use
Negative externalities of spectrum use	Spectrum and Health	Real Decreto 1066/2001	Protection of health against electromagnetic emissions
Competition Standardization	Access to common telecom infrastructure in buildings	Real Decreto-ley 1/1998 Real Decreto 346/2011 Orden ITC/1644/2011	Conditions for telecom infrastructure in buildings: - Access to the infrastructure - Standardization of technical features
Standardization-Enforcement of property rights	Standards for electromagnetic devices	Real Decreto 138/1989 Real Decreto 1580/2006	Definition of standards: -Electromagnetic compatibility

Source: Authors