Università
Ca'Foscari
Venezia

Master's Degree

in Management

Final Thesis

Streaming technology and its impact on the music industry

A Logistic Regression Analysis of Willingness to Pay for digital music and premium subscriptions to streaming music platforms

Supervisor
Prof. Daria Arkhipova

Graduand
Francesca Savoia
860338

Academic Year
2019/2020

Abstract

When it comes to digital transformation, one of the fastest and most radical revolutions ever happened, is that of the record industry. Part of this transformation took place thanks to the introduction of a digital technology, well-known today under the name of streaming. This research project aims to investigate how streaming as digital technology has upset and is still changing the music industry and the record market. Following an in-depth study of the previous literature on the subject, this study focuses on addressing the relationship between the spread of music streaming platforms and piracy, through illegal downloads on online peer-to-peer sites (P2P). The analysis is supported by the implementation of a survey that sought the willingness to pay for digital music, in particular for a subscription to streaming platforms. A sample of 849 people aged between 18 and 35 years, both Italian and foreign, was analyzed. The methodology used for the data analysis is based on an empirical analysis using a binary logistic regression model. This study, through the survey carried out and the subsequent analysis of the data collected, aims to establish a relationship among different variables that influence the willingness to pay (WTP) for a premium subscription to a music platform and the willingness to abandon (WTA) the use of P2P sites of illegal download leading to a greater understanding of the phenomenon of piracy in order to be able to reduce its use and spread, which has been threatening the entire record industry and the revenues of record labels and artists for centuries.

Table of Contents

List of Abbreviations

AI	Artificial Intelligence
AIDC	Automatic Identification and Data Capture
BaRT	Bayesian Additive Regression Trees
CD	Compact Disc
HD	High-Definition
HiFi	High Fidelity
IFPI	International Federation of the Phonographic Industry
IoT	Internet of Things
ISPs	Internet Service Providers
IT	Information Technology
LP	Long Play
MP3	MPEG Audio Layer-3
MPEG	Moving Picture Experts Group
MQA	Master Quality Authenticated
OLS	Ordinary Least Square
P2P	Peer-to-Peer
PC	Personal Computer
PR	Public Relations
Q1 (Q2, …)	Fiscal Quarter
R&D	Research and Development
RIAA	Recording Industry Association of America
RPM	Revolutions per minute
RTP	Real-Time Protocol
RTSP	Real-Time Streaming Protocol
SACEM	Société des Auteurs, Compositeurs et Éditeurs de Musique
WTA	Willingness to Accept
WTP	Willingness to Pay
WWW	World Wide Web

List of Figures

List of Graphs

List of Tables

Introduction

The music industry, and in general the entire creative industry, represents an environment that does not perfectly follow all the rules and dynamics of all other general industries, because by offering a consumer good of entertainment – enjoyment, the music industry has its own unique dynamics. For this reason, very often creative industries have encountered the problem and the difficulty to monetize their products and services. In particular in recent years, with the spread of digitization, musical goods have been distributed exponentially through the Internet and online platforms, not always respecting the constraints of copyright that protect the income of the rights' owner of the good, a phenomenon now widely known that takes the name of piracy. This violation of rights and unregulated distribution of music has led to great losses and damages not only to artists – musicians but also to all parties involved in the supply chain – i.e., the process of production, distribution, and sale of these products.

When streaming, as digital technology, was born it brought about a further change for the creative industry, especially to audio-video products and services. With the introduction of platforms and agreements with record companies, a new way of distributing and consuming music developed. However, the relationship between piracy and the advent of streaming, the influence of streaming on illegal downloads, and most importantly, whether there has been an improvement in creative industry revenues since the spread of these platforms has not yet been studied in detail. What is accessed today, both from statistical data and multiple studies conducted in the previous literature, is that there are unfortunately still many users who do not pay royalties to the authors of recorded music. The remuneration of artists, singers, and musicians, therefore, remains a largely unaddressed issue.

This study, born from a deep personal passion for the music industry, aims to determine whether there is a relationship between these two phenomena, piracy and music streaming technology and aims to identify what factors influence a user's willingness to pay for subscriptions to legal music streaming platforms. With the hope of shedding light on the dynamics of the creative music industry, the analysis conducted in this dissertation also seeks to make its own contribution to the previous literature on the topic.

This dissertation is divided into six macro chapters and introduction and conclusion sections. Chapter 1 summarizes the history of the music industry from the beginnings of recording to the present time and outlines a general picture of what the record market is to date by describing its major players, stakeholders, and significant percentages. Chapter 2 introduces and describes the digital streaming technology, explaining how it works and its application to music streaming

platforms. Chapter 3 describes the phenomenon of music piracy and in particular, it describes its adoption through P2P file-sharing platforms. The main illegal downloading platforms are presented, and it is highlighted the impact the piracy phenomenon has had to the detriment of the music industry over the years – since the birth of digitization. Chapters 1, 2, and 3, therefore, represent a general theoretical framework describing the main issues that affect the analysis completed in this study. Chapter 4 reviews the existing literature in the field and provides a theoretical framework for this analysis, several previous studies on the topic are presented highlighting their findings and possible limitations that need to be addressed. Chapter 5 presents the data collected for analysis through the survey, key summary statistics, and aggregate level data. Chapter 5 also describes the empirical model chosen for the study, the binary logistic regression model, and the variables included in the model, explaining their meaning, and describing their use. The main body of the thesis concludes with the analysis itself and the interpretation of the results in chapter 6. Lastly, sections following chapter 6 discuss and conclude the thesis, with comments on the practical limitations of this study and recommended future research.

CHAPTER 1. The music industry and the record market.

1.1. History of the Music Industry

Below a short journey through what has been the history of recorded music from the invention of the phonograph in 1877, which represented the first possibility people had to record and play music without having to resort to a live performance, to the present days with the advent of streaming technology and the birth of music streaming platforms.
The paragraph is divided into three sections which describe the pre-digital era, the digital era, and the streaming era respectively.

The Pre-Digital Era

The expression *pre-digital era* indicates the period between 1877 and the late 1980s – middle 1990s. The invention of the phonograph in 1877, represents the event that marks the beginning of the pre-digital era lasting until the 1980s and then ending with the subsequent invention and diffusion of the digital format MP3.
Thomas Alva Edison announced the invention of the phonograph one hundred and forty-four years ago, on November 27, 1877. The phonograph was a machine capable of recording and reproducing sound. Thanks to this instrument, for the first time listening to music, became accessible without the need for a live performance.
The development of sound recording began to function as a disruptive technology to the commercial interests which published sheet music. During the previous times, in the so-called *sheet music era*, if a person wanted to hear popular new songs, he or she needed to buy the sheet music and play it at home on the piano, which was the most used instrument of the years. Thanks to the introduction and spread of music recordings, consumers and insiders had the opportunity to listen to productions repeatedly without necessarily needing someone to read music from sheet music and play it with the appropriate instruments.
The idea of fixing sounds on physical support was a winning one and shortly after Edison's invention, one of his rivals, Emil Berliner, developed another successful machine, the gramophone. Emil Berliner invented the gramophone in 1887 and unlike Edison, he thought about conceiving support for music from the beginning, making possible the reproduction of sound in a large number of copies, the record.
It was not possible to duplicate Edison's cylinders for a long time, and for this reason, musicians were forced to record performances many and many times, and record production companies to

sell in the market recordings of different and inconstant quality, in insufficient quantity to a mass market. Berliner's records instead spread.

The concept of the *modern* music industry began during the time period between 1930 and 1950 and is marked by the concept of recordings (Pingitore, 2016).

The beginning of the diffusion of musical material on a large scale began in the 1940s with the advent of vinyl, in its 78, 45, and 33 rpm versions, in which the storage of sounds was done in analog through grooves on the surface of the record (Longo, 2014). While brands such as Fender and Marshall made their way in live reproduction, in terms of personal and domestic use it was the gramophone that marked the beginning of the reign of vinyl. Discs of various speeds and materials had been around since the beginning of 1900; the first versions rotate at 78 rpm, and the material used to produce them was shellac.

In 1948, Columbia Records produced the 12-inch 33 rpm *long play* format, the first LP ever released is titled ML4001 and is a *"Mendelssohn's Violin Concerto in MI Minor"* by violinist Milstein with the New York Philharmonic-Symphony Orchestra, conducted by Bruno Walter. Shortly after, RCA Records develops a 7-inch to 45 rpm *extended-play single* format, also known as EP. Due to the fragility of shellac, which often breaks during transport, Columbia and RCA Records began producing their EPs and LPs on vinyl (Donà, 2017). The big flaws of vinyl were size and portability. To find a solution to these problems, the industry decided it had to develop new formats that people can easily take with them to work, parties, etc.

The first change occurred in the 1960s when the vinyl has then been substituted by the cassette tapes and the debut of the Walkman, the portable audiotape player introduced by Sony. The introduction of these devices responded to the need to be able to listen to recorded music also outside the home. While holding the same amount of information, a blank cassette was sold for around $3 against the $6 of the vinyl. Compact cassettes or tapes were invented by Philips and introduced in Europe at the Berlin Radio Show: the oldest technology convention in Europe. The first cassettes featured *reverse housing* with a maximum time of 45 minutes of stereo audio per side, significantly longer than the vinyl LP playback time. Tapes, on the other hand, were packaged more conveniently and compactly. The small size of the tapes gave rise to portable players, making them a convenient development in the history of how and where people listen to music. Indeed, the success of the cassette tapes was in the reduced price, the small size, and the possibility of recording events shows, or radio programs. The cassettes also fit perfectly into the post-war era characterized by a boom in population and suburban expansion, meaning an increase in the number of cars and consequently an increase in the demand for mobile playback systems and formats. In this period the first record production companies were born, which represented the way for artists to achieve

success. The record industry replaced the sheet music publishers as the music industry's largest force. The first record production company was founded in 1888 in New York and bore the name Columbia Records. Originally the company dealt exclusively with the sale and distribution of phonographs and phonographic cylinders but, in a few years, the business grew first with the production of its media and, shortly after, with the introduction of the first records, first the 'XP' and then the 78 rpm. From 1912 Columbia dedicated itself exclusively to recording on disc and today continues to carry on its activity as a subsidiary label of Sony BMG Music Entertainment.

Many labels were born, some still exist today, and others died quickly, among the best known there are Crystalate, Decca Records, Edison Bell, The Gramophone Company, Invicta, Kalliope, Pathé, Victor Talking Machine Company, and many others.

During the period when cassettes were gaining popularity, another vinyl competitor was being born, the 8-track. The 8-track was a collaborative invention of the three companies: RCA Records, Lear Jet Company, and Ampex Magnetic Tape Company (Donà, 2017). The advantage of 8-track tapes over the compact cassette tape is their ability to accommodate 8 parallel audios with four matching stereo programs, i.e., they could play a lot of music in a relatively small package. In fact, with its introduction, in the middle 1960s, consumers were able to record twice as much music as they could with cassettes. Qualities that also made the 8-track a valid vinyl rival were its portability and more accessible use, for these reasons in 1974, the 8-track became the fastest-growing format in the industry. Much of the success of the 8-track tape is also related to the explosion of the automotive industry at the time. As early as 1966, Ford Motors offered 8-track tape players as an option to be installed in their complete line of cars produced that year. Despite the popularity of the 8-track in the 60s and 70s, the compact cassette tape becomes the most popular choice for artists and consumers for reasons of price and size. The last 8-track tape released by a major is said to be Fleetwood Mac's Greatest Hits, released in November 1988 by Warner Records.

During the 1980s and 1990s, some artists started publishing albums in the floppy disk format, even though they are normally associated with data storage for desktop computers. IBM introduced the 8-inch floppy disk to the technological world in 1972, followed in 1976 by a 5-inch and ¼-inch model, later replacing it with an even smaller 3-inch and ½-inch format in 1982. This format remained quite a niche and did not really become mainstream. The most famous floppy disk is the album by *Brian Eno Generative Music I,* released on Opal Music in 1996 (Donà, 2017). Regardless of the difficulty of the floppy disk to break into the music market, the floppy disk represented an important vision for the digital future of music; a trend that was soon resumed by the explosion of another important format: the CD – Compact Disc.

A relevant change is registered with the arrival of the 1980s, which were the watershed between the two great technological eras: that of *analog* and that of *digital*. With the introduction of the new digital support, the CD, and personal computers, a profound transformation of the industry took place in a very short time. In 1980 the music tracks in computerized format radically changed the rules of the game. In fact, the first music album on CD launched on the market, 52nd Street by Billy Joel, dates back to 1978. The compact disc was first used by the music industry in 1982 and soon spread around, it seemed to bring together the best of all the formats that came before it: high sound quality, compact format, portability, ease of use, re-writability, and inexpensive audio. The CD was an extremely important development for the music industry, becoming the standard release format for decades, replacing any other musical support, and therefore becoming the main source of revenue for record companies. In a sense, CDs signed the beginning of the end for physical formats. The industry quickly adapted to the demands of the public by launching one format rather than another depending on the needs so as never to lose consumer confidence and market control. The consumer, thanks to the development of new technologies, slowly began to take possession of some tools that until then were for the exclusive use of professionals; from burners to file sharing, which definitively knocked out the music industry in a very short time. In 1987, sales of records from CD finally surpassed those of vinyl, which were in decline. This proves that consumers' power influences the way music is recorded and sold and therefore represents a major innovation in the music industry.

As mentioned above, in the late 1980s, many record companies were born but many also quickly died, those that survived and dominated the industry were known as the *Big Six* – EMI Electric and Musical Industries, CBS (now Sony Music Entertainment), BMG Bertelsmann Music Group, PolyGram, WEA Warner Music Group Corp. and MCA which later gave way to Universal Music Group.

The Digital Era

The 1990s indicate a breaking point that defines the beginning of the *record industry* and it is precisely marked by the birth and subsequent spread of record labels. Moreover, the 1990s is a revolutionary period of innovation for the music industry, as it brought about the relevance of music in its digital form. The advent of the internet and digitalization has characterized the first decade of the 21st century with a devastating impact on the markets and the sales of major record labels. The combination of the internet and digital audio recording made possible the birth of the MP3 format. Technological progress in 1995, realized the potential of a new format in the context of the spread of the Internet, the .mp3 extension. MP3, formally Moving Picture Expert Group-1/2 Audio Layer

3, is a lossy audio compression algorithm, developed by the MPEG group, capable of reducing the amount of data required to store a sound, while maintaining a faithful reproduction of the original uncompressed file. MP3 allowed consumers to enclose audio files in a space ten times smaller than that used in a CD audio so that they can be transferred via the Internet.

Its conception was due to a work team established at CSELT and coordinated by the Italian engineer Leonardo Chiariglione (Chiariglione, 1998). The implications of this innovation for the music industry took on enormous and unforeseen proportions and shaped forever the performance of the music industry and the behavior and interests of its consumers.

In the 2000s the music market was controlled by the three so-called *majors* record labels: the French Universal Music Group, the Japanese Sony Music Entertainment (Sony, 2008), and the US Warner Music Group. The record labels outside of these three formed the group of *independent* labels or *indies* (Pingitore, 2016).

During this period there has been another very important change in the music industry, marked by the spread of the Internet and therefore the spread of illegal file-sharing sites P2P peer-to-peer. The popularity gained by illegal P2P sites has led to the birth of the well-known problem that characterizes the music industry for years: piracy. In 1999 two young American computer scientists, Shawn Fanning and Sean Parker, with the introduction of Napster, proposed the idea of sharing entire music repertoires among users of a network in a completely free way by bypassing the law and copyright and actually opening the door to digital piracy (Grassini, 2012). This program allows people to share entire portions of their hard disk, containing MP3 files and it is identified with the name of P2P. P2P stands for peer to peer, in telecommunications indicates a model of logical computer network architecture in which nodes are not only hierarchized in the form of fixed clients or servers, but also in the form of equivalent or equal nodes (peer), being able to act at the same time as clients and servers to the other terminal nodes (hosts) of the network. Through this configuration, any node is able to start or complete a transaction. Equivalent nodes may differ in the local configuration, processing speed, bandwidth, and amount of stored data (Benayoune and Lancieri, 2004).

The P2P file-sharing phenomenon has caused not only a social transformation, as the exchange of songs online becomes an experience and a possibility of fun through the sharing of interests among end-users, but at the same time has allowed the record department to make digital distribution faster and more convenient than traditional distribution.

The functioning of Napster involved the connection to a central server which was able to read all the files of the connected computer that had the .mp3 extension; this server was interrogated by other users who could search for the desired track and download it directly from the hard disks of

other connected users (Bartsch, 2017). In just two years, millions of files were shared, i.e., this also meant millions of copyright violations. Napster gave users the possibility to record CDs into their personal computers and to make tracks available for download to everyone simply logging in to the Napster network. To understand the popularity and the impressive impact Napster had on the music industry it is possible to simply look at the fast and enormous growth it achieved without any form of marketing; it was estimated that Napster had seventy-five million registered users downloading approximately ten thousand songs per second (Ku, 2002). Because of the huge copyright infringements and economic losses that Napster was causing, RIAA – *Recording Industry Association of America* – the association that groups together the most important American record companies, officially sued Napster in the court of San Francisco in 1999. In September 2001, the American court, the *"United States Court of Appeals for the Ninth Circuit"* at the end of the trial that involved Napster, Inc. against the A&M records label, ordered the immediate cessation of the activity carried out by Napster as harmful to copyright protection, imposing compensation to record labels for $26 million (Longo, 2014). The active users at the time of the process were about 20 million with forecasts of a dizzying increase to 70 million. Napster was not the only illegal site but was the first of many other peer-to-peer programs such as eMule, WinMX, and BitTorrent, born to exchange multimedia (Peron, 2016).

Because of the free access to music, the great problem of music piracy was developed not only by illegal download sites such as Napster but also through different models of file sharing, such as social networking sites or music video platforms. This phenomenon, that so much frightened the big music multinationals – majors, however, showed the overwhelming interest of the mass of users to appropriate digital music files, leading market strategies to the creation of suitable support for the enjoyment of digital music: a new mp3 player: the iPod.

Steve Jobs, the creator of Apple Inc., launched the iPod in October 2001 and in the same year released iTunes, a response to the Napster phenomenon. iTunes was firstly available in 1998 under the name of SoundJam MP but was released just in 2003 with the name it has today, iTunes Store (Longo, 2014). Initially, SoundJam MP lacked popularity due to Napster, users in fact enjoyed the free downloads on the illegal site. iTunes Store worked together with the innovative iPod products, and for this reason, Apple dominated the digital market for the rest of the decade (Pittman, 2016).

Parallel to the iPod generation, YouTube, a totally unexpected rival in the music industry, was gaining popularity; a video broad-casting site created in 2005 with the intention of allowing the diffusion of recorded videos of its users. Soon it was used as a medium for spreading counterfeit music videos. The channel lent itself perfectly to the publication of videos of all kinds in a completely free and almost uncontrolled way. This is where music streaming first came in.

YouTube offered users the possibility to view video clips of artists from all over the world without the need to download them to one person's PC. In 2006 YouTube was acquired by Google for $1.65 billion in its own shares and today YouTube is the second most visited site worldwide after Google.

The Streaming Era

The middle 2000s introduced an interesting concept of music consumption: the streaming technology (Fly, 2016). With 24/7 Internet access and an increase in technological familiarity, developers and entrepreneurs saw the opportunity to deliver music in a new and innovative way, allowing users to gain the ability to listen and discover new music without actually having to download files or buy tracks to a computer or device. Streaming applications filled the growing demand for non-physical access to music and open the most current chapter on formats: dematerialized music. This new concept overcomes the idea of *liquid music* and familiarizes with the meaning of digital music available through the access to a virtual cloud, a large library containing billions of tracks and albums at everyone's disposal, at any time, simultaneously, through any device and in any place.

Through the streaming technology, the music was simply delivered as a continuous stream of data. Streaming platforms aimed to make digital music a sustainable business model that could bring benefits for all players involved. The music streaming platforms aim to create a *win-win* strategy for everyone involved in the process; the majors have revenue from listening to music files, users can listen to music in a legal way and at affordable prices, musicians have the opportunity to reach a very large audience, and distributors have ensured the development of a legal model to achieve profits (Longo, 2014). The Swedish company Spotify was founded by Daniel EK and Martin Lorentzon, which conceived the first streaming subscription service, with possible insertion of advertising and limitations of use in case the user wants to use it completely free of charge. Spotify was launched in October 2008 and according to Spotify's Q1 2020 report, there were 286 million monthly active users of which 130 million were Spotify Premium subscribers. The business model offers users two choices: listen for free with the interruption of ads or pay a monthly fixed fee for unlimited and uninterrupted streaming. According to RIAA data, the definitive overtaking of digital revenue from the physical ones occurred in 2014, in the years after, the gap that has been created between streaming and physical has gradually grown.

Given the recent events caused by the COVID-19 pandemic that has affected the whole world, some music platforms are innovating by hosting virtual concerts for artists. As explained in the article published by Deloitte Insights of Srivastava, S. and Downs, K., even though *"it may be*

impossible to re-create the concert experience in a living room, many digital platforms are bringing artists closer to fans", e.g., Tidal offered its clients free access to virtual. Cercle is an example of a live stream media company dedicated to promoting artists. Every Monday Cercle offers its users the opportunity to watch a 1-hour concert in breathtaking locations broadcast live on the company's Facebook page. After the performance, users can ask questions online to the guest who just performed. Another epic example is last year Travis Scott's free virtual concert inside Games' Fortnite that reached 12 million viewers.

"Streaming platforms have witnessed increased social engagement, subscriptions, and artist activity through their COVID-19 innovations" (Downs and Srivastava, 2020).

1.2. The Record Market

Analyzing the economic context, the music market is defined as a vertically differentiated oligopoly, the entire sector is dominated almost entirely by a small number of companies, the *majors*, taking over almost the total market share. The majors set the prices and the market presents high barriers to entry. The market can be defined differentiated as the quality, packaging, and branding of music companies would be unique, it focuses on the sale of record products heterogeneous between them (Grassini, 2012).

The majors are characterized by a structure that presents a macro division that deals with the work on the musical product. This is divided into four levels: artistic level (1) dealing with the development of the musical product, marketing and promotion level (2) concerning the advertising of the artistic product on the market and all available tactics, techniques and strategies, publishing level (3) characterized by the copyright management of the product and distribution level (4) representing the commercialization on the market of the artwork. The independent labels, indies, are generally more specialized in niche markets dealing with specific genres (Gramolini, 2015).

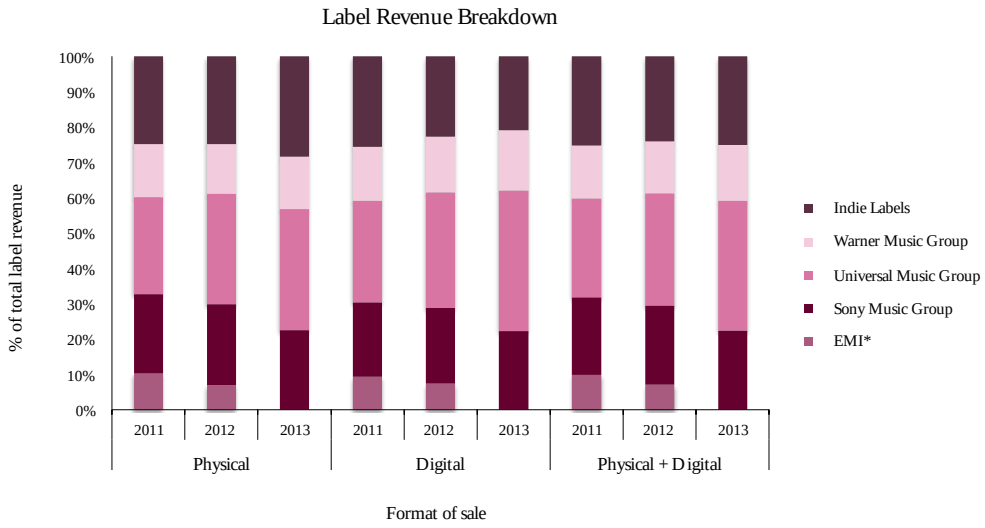
As mentioned above, the market is controlled mainly by the majors, which owned more than 67.5% of the total market share in 2019 – half-point below the 68% of 2018 – they are known under the name of *Big Three*: Sony BMG, Universal Music Group, and Warner Music Group.

Once known as the *Big Five*, in 2004, following Sony Music's acquisition of Bertelsmann Music Group, better known as BMG, they became the *Big Four* and since 2012, after further EMI absorption by Sony and Universal, they have settled into the *Big Three* (Boccanegra, 2015). The UK record company EMI Electric and Musical Industries was the notorious label of stars such as The Beatles, Coldplay, Katy Perry, Kylie Minogue, Lily Allen, Norah Jones, and Robbie Williams. In 2012, Universal Music Group, the wholly owned subsidiary of the French group Vivendi, bought

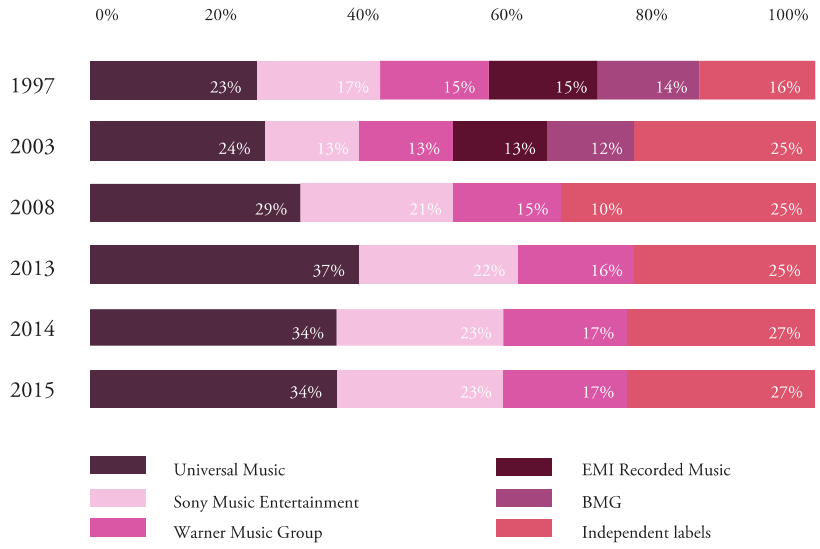
the UK record company Emi from the US bank Citigroup for £1.2 billion, approximately €1.4 billion.

The majors retained the share of the overall market in 2019, accounting for 67.5% of the total while independent labels and direct artists, i.e., those who are without a record label, accounted for the remaining 32.5% of the total market share, it is considered the fastest-growing segment in the market, reaching a value of $873 million, which is equivalent to 4.1% of the total (Mulligan, 2019). Graph 1 shows the labels revenue breakdown – physical, digital, and overall – during the period 2011-2013, before the acquisition of EMI.

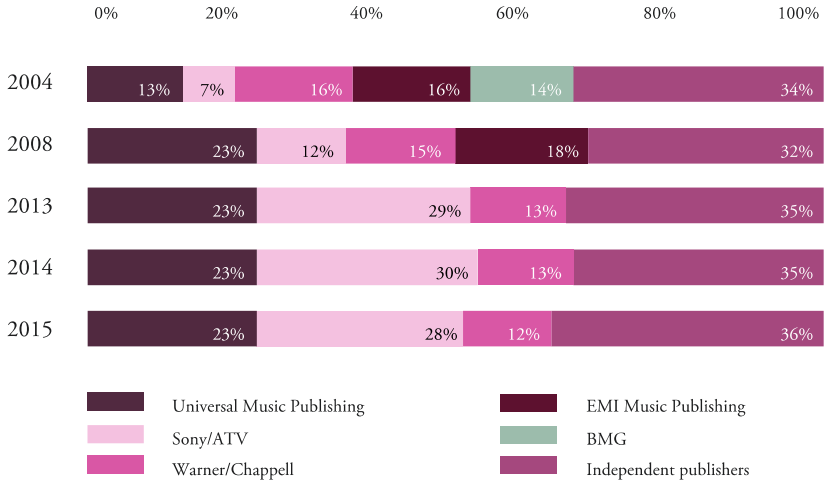
Graph 1. Label Revenue Breakdown (RIAA, 2020).



Graph 2. Global Record Industry Market Share (1997-2015).



Graph 3. Global Music Publishing Market Share (2004-2015).



In order to take a broad and comprehensive look at the present global music market situation, the data below is based on the *Global Music Report 2020* issued by the *IFPI – International Federation of the Phonographic Industry*, reporting all music market statistics at the end of the year 2019. As specified in the report, the analysis was made before the advent of the world pandemic by COVID-19. For this reason, forecasts and growth data will certainly be slightly distorted because, as ascertained, the pandemic has brought a huge change in the economy and between the sector and the industries most affected there are the creative ones and in particular, the music, events, live and performance sectors.

Analyzing the numbers and data, the report clearly shows positive growth in global revenue of 8.2% in 2018 reaching the value of US$20.2 billion, celebrating the market's fifth consecutive year of growth. The data relative to streaming is a growth of 22.9% in overall revenues deriving from streaming, a growth of 24.1% in paid streaming revenues, and an important growth in the percentage of paid streaming subscribers of 33.5%, the number of paid streaming accounts raised to 341 million by the end of 2019. The percentage in shares, streaming has over global revenues is 56.1%, exceeding more than half of the total, compared to all other forms. However, this general growth in streaming and subscriptions has led to a decline in revenue from physical of 5.3%.

As analyzed by the Global Music Record *"The world's top ten markets all demonstrated growth, with the exception of Japan. There was particularly strong growth in paid streaming with all top 10 markets experiencing double-digit growth in paid streaming revenues"* (Moore, 2020).

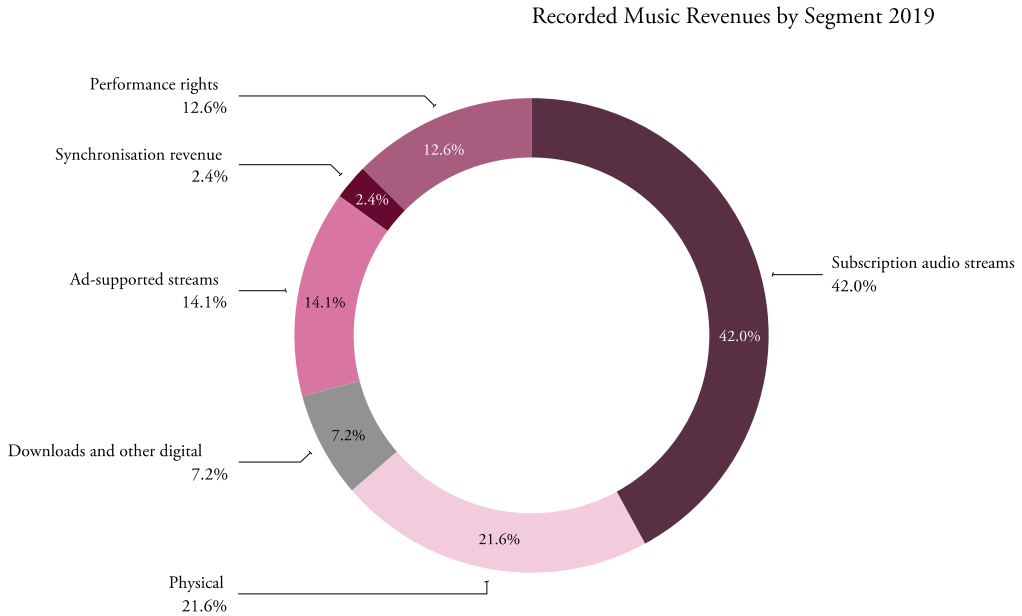
Graph 2 shows the global recorded music revenues by segment in 2019; looking at the revenues divided by different formats, the report demonstrates that streaming is the dominant format

globally, overall streaming revenues climbed by 22.9% in 2019 to US$11.4 billion. Digital revenues, from downloads and other digital – streaming excluded – experienced a decline of 15.3% in 2019. It has been hypothesized that this drop is due to the steep decline across multiple markets in download revenues, which accounts for just 5.9% of the total. Latin America was the region with the steepest decline in non-streaming digital revenues (RIAA, 2019). Revenues from performance rights posted a decline of 3.6% in 2019 with respect to the previous year but showed a comparable growth rate of 8.7% over 2017. Talking about physical revenues, in 2019 it has been experienced a decline of 5.3% globally, but growth in a small number of countries e.g., the USA and Spain.

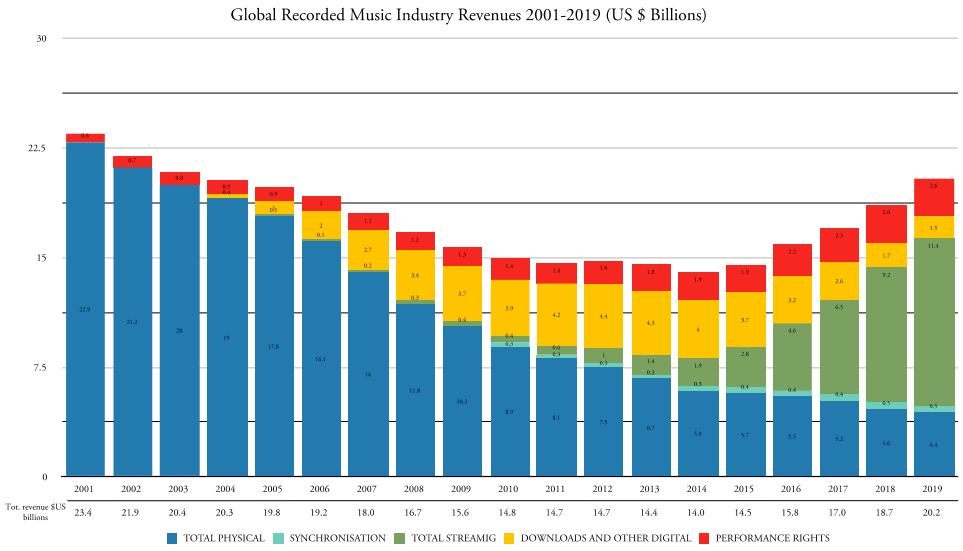
In particular, revenue from vinyl increased by 5.3%, represented today 16.4% of overall physical revenues. Synchronization revenues, the ones deriving from the use of music in advertising, film, games, and TV experienced a growth of 5.8%.

Analyzing the different markets by regions; the USA and Canada remains the largest region for recorded music revenues, 39.1% of the global market, Europe represents the world's second-largest region and experienced growth in some of the region's biggest markets like the UK, Germany, Italy, and Spain (Moore, 2020).

Graph 4. Global Recorded Music Revenues by Segment 2019.

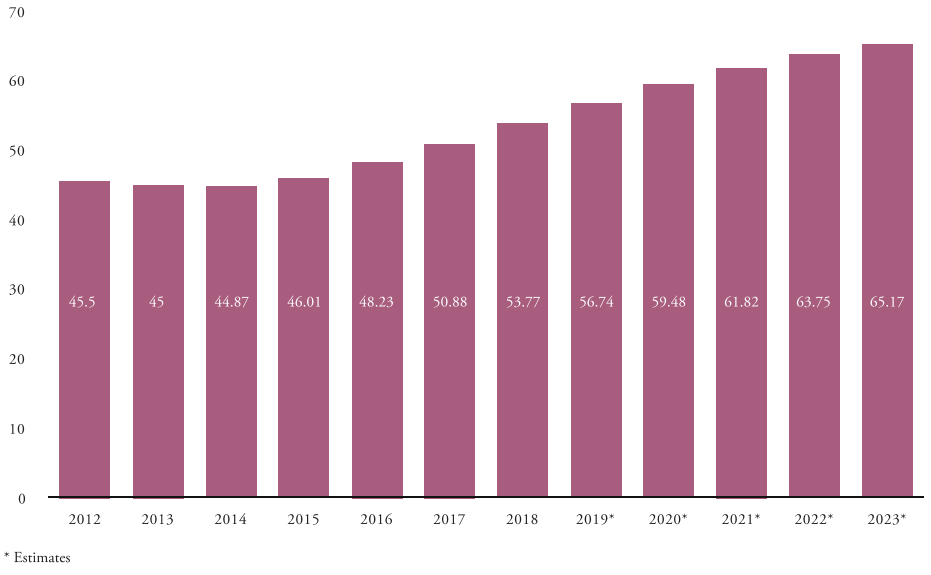


Graph 5. Global Recorded Music Industry Revenues 2001-2019 (US $ Billions).



Graph 3 displays the global revenues for 2001-2019 of the Recorded Music Industry (Moore, 2020). As shown, in 2015, digital revenue officially exceeds physical revenue by 45% compared to 39%, and the sector is experiencing its first recovery after several years of decline. This trend continues in 2016 with the market up 5.9%, as evidenced by the IFPI report in April 2017, and was the highest growth rate since IFPI began to follow the market in 1997 (Leocata, 2017).

Graph 6. Music industry worldwide from 2012 to 2023 (in billion U.S. dollars)[1].



[1] PwC; Ovum; IQ Magazine; Statista 2020

Below are shown the *RIAA* statistics – *Recording Industry Association of America*, a trade organization that represents the recording industry situation regarding the United States.
For the fourth consecutive year, the U.S. record market saw double-digit revenue growth of 13% in 2019 from $9.8 billion to $11.1 billion. There has been a steady increase in subscriptions for paid services, now boasting more than 60 million subscribers.
A closer look at streaming shows that *"total revenues from* ***streaming*** *grew 19.9% to $8.8 billion in 2019, accounting for 79.5% of all recorded music revenues"*. When talking about the streaming world, RIAA specifies that in its analysis *"the streaming category includes a wide variety of formats, including premium subscription services, ad-supported on-demand services (such as YouTube, Vevo, and ad-supported Spotify), and streaming radio services (like Pandora, SiriusXM, and other Internet radio services)"* (RIAA, 2019). The streaming market alone in 2019 was larger than the entire U.S. recorded music market just 2 years ago in 2017.
Paid subscriptions to on-demand streaming services contributed by far the largest share of revenues and accounted for 61% of the total, with a value of $6.8 billion. On-demand streaming services supported by advertising streamed more than 500 billion songs to more than 100 million users in 2019 in the USA.
As specified by the RIAA statistics, the formats that experienced a collapse in total revenue were the digital formats, other than streaming, such as digital downloads which decreased by 18% to $856 million in 2019, the first time in 14 years that revenues from downloaded tracks and albums fell below $1 billion.

1.3. Music Industry Stakeholders

The stakeholders that characterize the music industry are creative artists such as composers, songwriters and musical performers, agents such as managers, promoters, etc., music publishers, record companies, copyright collecting, studio owners, manufacturers, distributors, retailers, broadcasters, venue operators, ticket agents, filmmakers, multi-media producers, advertisers, and finally, individual consumers, who purchase a musical good or service (Ahmadi 2015) (Throsby, 2002).
In the analysis carried out by Ricardo Álvarez Vázquez *"The music industry in the dawn of the 21st century Networking for a thriving music industry"* different groups of stakeholders are identified. The models divide the industry into different subcategories:

- the *publishing industry* – involving the acquisition of rights and licenses –is characterized by actors such as composers, songwriters, and lyricists, music publishers, authors' collective management organizations, licensing agencies and sync representatives, sheet music

publishers, online licensing – i.e., sync representatives – and royalty collection services, copyleft and creative commons licenses, and online lyrics aggregators and publishers.

- the *recording industry* – meaning the part of the industry involved in the discovery and development of talent to produce albums and in the promotion of artists and their recordings. The parties related to the recording industry for this model are: performers, performers' collective management organizations, managers, music producers, recording studios, mastering studios, record labels, physical media manufacturers, music distributors, physical retailers, digital music distributors and aggregators, non-interactive streaming music services – i.e., ad-based and subscription-based – on-demand or à la carte streaming music services – ad-based and subscription-based – pay-per-download stores, music lockers, hybrid music services, video streaming services, video sharing services, video games manufacturers, background music service providers, label services, online music databases, podcasts, ringtones, P2P networks, home audio equipment manufacturers.
- the *live music industry* – including concerts promotions and productions – is characterized by stakeholders such as booking agencies, event promoters, festivals, ticket sales, and distribution companies, music venues, live music, and event streaming, concert listings, online ticket sellers, and resellers – or tickets exchange.
- other players – stakeholders that cannot be directly associated to any of the three main subcategories like automatic content recognition and rights monitoring, automatic identification and data capture (AIDC) technology, blockchain technology, cryptocurrencies, smart contracts, brand and advertising, R&D specialists, consumers, cross-media platforms, crowdfunding platforms, educational institutions, film industry, financial and bookkeeping services, government agencies, instant messaging services, legal counseling services, merchandise manufacturers and retailers, musical instruments manufacturers, music awards and music contests, music blogs, music charts, musicians' unions and guilds, music intelligence and analytics, music promotion platforms, online communities and forums, online – mobile and P2P payment services, PR agencies, printed media, specialized magazines, fanzines, professional audio manufacturers, prosumers, netlabels, online mastering services, radio, rehearsal spaces, review aggregators, social networks, trade fairs and conferences, trade associations, trademark offices, TV, webzines, dancers and choreographers (Álvarez Vázquez, 2017).

CHAPTER 2. The Streaming Technology

2.1. Streaming as a Digital Technology Innovation

Clayton Christensen in the early 1990s coined the term *disruptive innovation* describing *"a process by which a product or service initially takes root in simple applications at the bottom of a market—typically by being less expensive and more accessible and then relentlessly moves upmarket, eventually displacing established competitors"* (Dillon, 2020). With the term *creative destruction,* it is intended a process of industrial change that incessantly revolutionizes the economic structure from the inside, destroying the old one and creating a new one. Innovation is part of a process and not a single event; when an organization fails to keep up with the revolutionary change, it will eventually be left behind and forced to leave the sector (Longo, 2014). Disruptive innovations are characterized by completely disrupting the previous market and streaming is an example because of its impact on music CD sales; its characteristics won over the existing market.

Streaming began to gain popularity in the early 1900s and has undergone a process of revolution and improvement since then. Streaming is defined as *"a method of transmitting or receiving data – especially video and audio material – over a computer network as a steady, continuous flow, allowing playback to start while the rest of the data is still being received"* (Oxford English Dictionary, 2020). Streaming allows users to access online content – e.g., audio, video, songs, movies – without having to download or possess them. This technology offers maximum convenience in accessing multimedia content; users with a reliable Internet connection will be able to play from their device an entire movie or album smoothly without interruption. In order for companies such as Netflix or Spotify to provide content for streaming, they need servers or cloud platforms for storage. They have content distribution networks that keep the most popular content cached and close to where it will be streamed to reduce latency and bandwidth costs. When streaming content, the data is delivered to the buffer, which stores the next few seconds or minutes of the song or TV program people desire to access.

Erroneously people often tend to associate the streaming technology with the download, unlike those who think they are two similar things, it is very important to understand that the two technologies are actually two very different processes. When downloading a movie or song, the file is saved to the user's hard drive. It is possible to start watching or listening to it only after the download is complete. Moreover, once a file is downloaded, it will be stored on the user's device and it will take up space on the hard disk. Streaming, on the other hand, plays multimedia content without downloading the files, so users will not consume valuable hard drive space. Streaming

allows the consumption of music without physical possession of the music file, albums and tracks are stored on a server and can be listened to on-demand by consumers (Delikan, 2010).

2.2. Different types of streaming and main players

In today's media landscape there is a wide range of content that can be streamed; music and videos are the most traditional formats, but more and more new streaming options are available including games, apps, and even live events. The companies which are the main players in the streaming sector are music platforms such as Apple Music, Pandora, Deezer, and Spotify and movie and TV program archives such as Netflix, HBO GO, Hulu, and Disney+. Although they are all classified as streaming platforms – since the service they offer is the same – they can be divided into different categories according to the type of streaming content they offer to their consumers:

(1) Streaming audio files, that include music and podcasts, allow people to play an infinity of songs from different albums, or artists, without having to download a single file. Services like Apple Music, Spotify, and Pandora make millions of tracks available for streaming at the touch of a button. Some, like Pandora, let people choose a genre or mood and then offer playlists tailored to their taste. Others, like Spotify, reproduce exactly what was selected, although playlists still play an important role in the platform. Apple Music offers a combination of both options. Podcasts can be streamed or downloaded for later listening and are available through services such as iTunes and Stitcher.

(2) Videos were the first content to be mass streamed, starting with the spread of services such as YouTube. Instead of having to download large media files, streaming video involves compressing the data into small packets that are sent to the user's device where they are decompressed and displayed. During a streaming session, the video continues buffering: while viewing one data packet, the next one is being decompressed, so it can be watched entirely without interruption. The most popular video streaming services are YouTube, Netflix, Amazon Prime, Hulu, Google Play, and Disney+.

(3) The streaming of games, or apps, works in the same way as audio and video. It saves space and limits processing requests on devices by uploading everything to an external game server.

(4) Live streaming works similarly to other types of content but is used specifically for special events such as sporting events or political debates. When watching a live stream, people visit a website (e.g. news) hosted on a web server. This connects to a media server, which transmits content to their device using a real-time protocol (RTP) and a real-time streaming

protocol (RTSP). In this way, the video files are sent in a smaller (compressed) format and then displayed in a higher quality (uncompressed) format on your device.

Recently, also due to nationals lockdown caused by pandemic COVID-19, social media platforms have started incorporating live streaming elements, such as Facebook Live and Instagram Live, offering users the possibility to access concerts or conferences from home.

One of the qualities that has made streaming so popular is convenience; users no longer must wait for content to download but simply start listening or watching what they want almost immediately. Moreover, another advantage of streaming is the cost reduction and an alternative to piracy. It could be expensive to buy many CDs or download digitally many albums of all the bands you like and downloading albums from P2P sites is illegal. Streaming gained popularity very fast because it offers users an alternative solution to both problems: by paying a monthly fee for a streaming service – usually around $9, the price of one physical album, people can avoid buying or illegally obtaining all the content they desire. Moreover, another feature of streaming platforms is their ability to let users have unlimited storage on the online "library"; unlike downloading, streaming will not store large files on devices, giving users the possibility to save songs, movies, or playlist without consuming storage on their computers or mobile phones. Recently, live streaming has become very popular, allowing users to watch events, conferences, sports games, or concerts directly from their devices anywhere.

2.3. Music Streaming Platforms

Spotify AB (a) is the global leader in music streaming, it offers on-demand streaming of a selection of songs from various record companies and independent labels, including Sony, EMI, Warner Music Group, and Universal. Spotify was developed starting in 2006 by *Spotify AB* in Stockholm, Sweden. The company was founded by Daniel Ek, former CTO of Stardoll, and Martin Lorentzon, co-founder of TradeDoubler, and was launched in October 2008. *"With Spotify, it's easy to find the right music or podcast for every moment – on your phone, your computer, your tablet and more. […] Choose what you want to listen to, or let Spotify surprise you"* (Spotify, 2020).

Spotify operates under a freemium business model and generates revenues by both selling premium subscriptions to users and advertising placements to third parties. Since its launch in 2008, Spotify's largest expense has been royalties, accounting for about $9 billion. The company has been heavily criticized as it once ranked as one of the industry's worst royalty payers, and for this reason, musicians and labels raised many complaints. This has even led some artists such as Taylor Swift and Radiohead to remove their music from the platform as a protest to the low royalties' payment

fees. Following these complaints Spotify is now steadily increasing its payments; its per-play rate was 0.00437 cents in 2019, according to Digital Music News. Spotify pays artists and labels based on their market share – meaning the number of streams for their songs as a proportion of the number of total songs on the platform. The company distributes about 70% to those who hold the rights, generally labels, who then pay the artists based on individual agreements made in advance.
Spotify's Q3 2020 report, released at the end of October 2020, revealed that the world's most popular music streaming service can count, as of the end of September 2020, on 144 million premium users, i.e., subscribers who pay to have access to all the features it offers. Spotify has also surpassed 320 million monthly active users globally, who use the free and paying services combined. Spotify is available in free and paid versions in most countries in Europe, almost all of the Americas, Australia, New Zealand, and some countries in Asia. The service can be accessed through Microsoft Windows, macOS, GNU/Linux, Google Chrome OS, Telia Digital-tv and mobile devices equipped with iOS (iPod/iPhone/iPad), Android, BlackBerry (in limited beta), Windows Mobile, Windows Phone, S60 (Symbian), webOS, Squeezebox, Boxee, Sonos, PlayStation 4, Xbox One, WD TV and MeeGo.
Music on the platform can be viewed by artist, album, label, genre, or playlist as well as through direct searches. Spotify provides two types of versions: *free*, which is available at the time of subscription to the service by logging in with a Facebook or Spotify account; the user can listen to an unlimited amount of music, but overlaid with visual and radio-like advertising and with the possibility of switching from one track to another in the playlist six times in an hour; and *premium*, which allows users to listen to music without commercial interruptions and access additional features such as streaming at higher bitrates, offline access to music, and mobile apps.

Pandora Media, Inc. (b), Pandora Internet Radio is one of the most popular music streaming services in the United States (Longo, 2014). Pandora was launched in 2000 in the United States as an online radio service that relies on the Music Genome Project, it allows users to discover new music according to personal tastes through the subscription of a paid contract.
Like Spotify, Pandora also offers a free service and a paid one without advertising. By registering, the site allows users to create virtual radio stations by entering a song or an artist the user likes. The system will then use an algorithm specially created by the Music Genome Project to search for songs similar to the one reported by the user, and then play the music that the listener may like. Normally radio stations and record shops group together track by genre, by collaborative filtering or by rating, contrary Pandora, organizes music by musical traits or genes; some examples could be

the gender of the lead vocalist, the tempo of the chorus, the level of distortion on the electric guitar, the type of background vocals.
Pandora's revenue comes from advertisements placed in playlists, between songs, making it make money in the same way radio does. The royalty payments also make up the largest operating expense of the platform. Pandora's platform is considered a leader in digital music streaming services and its monthly active users total 63.5 million in 2019.

Apple Music (c) is a music streaming service developed by *Apple Inc.* It was revealed on June 8th, 2015 at the *Apple Worldwide Developers Conference* and released on June 30th of the same year, with the iOS 8.4 update, in about a hundred countries around the world. Apple Music is a service that *"allows users to listen to 70 million songs, ad-free, to download music and listen to it offline, see the lyrics flow along with the songs, use your library on all your devices and always discover new things thanks to personalized recommendations and playlists curated by the platform's editorial staff"* (Apple Music, 2020).
The platform was originated thanks to the purchase, for over three billion dollars, of the company *Beats Electronics*, a company that produces medium-high quality headphones for music reproduction. Actually, Apple's interest was not based on this Beats product, but mainly on the music streaming service that had been created by the company (Longo, 2014). Apple Music, unlike Pandora and Spotify, presents only one way of fruition, that is through a monthly subscription, following a three-month free trial. There are therefore no advertisements or limitations of any kind, and it boasts the entire catalog of music present on Itunes. Today Apple Music is present in 167 countries (Billboard, 2020). What makes it peculiar is the presence of the radio station *Beats 1* that broadcasts from New York, Los Angeles, and London 24 hours a day exclusively for Apple. The playlists are not created with special software but by experts and musicians who propose them for every moment of the day. An additional feature is the inclusion of music videos, a service that Spotify does not offer (Leocata, 2017).

TIDAL (d) *"is an artist-owned global music streaming and entertainment platform that brings artists and fans closer together through unique original content and exclusive experiences"* (Tidal, 2021). Tidal was launched in 2014 by Norwegian publicly held company Aspiro and consequently purchased by American rapper and business mogul Jay Z in 2015. The artist-owners of the platform are Alicia Keys, Arcade Fire (Win Butler and Regine Chassagne), Beyoncé, Calvin Harris, Claudia Leitte, Clifford "T.I." Harris, Coldplay, Daft Punk, Deadmau5, Jack White, Jason Aldean, J. Cole, Kanye West, Madonna, Nicki Minaj, Rihanna, Shawn "JAY Z" Carter, Damian Marley, Indochine, Lil Wayne, and Usher. TIDAL's commitment is to implement a different and innovative model for the music industry,

sustainable and fair towards artists and their rights. TIDAL generates revenue through the payment of a subscription, so it does not offer, as has been observed for other platforms, the free mode (TIDAL does, however, offer the possibility of trying a 30-day free trial). This platform offers two different subscriptions, *premium,* and *HiFi*; in the HiFi subscription, in addition to the features and characteristics already present in the premium one, music is provided in lossless quality, CD and Master Quality Authenticated (MQA) (1411 kbps vs. 320 kbps for standard streaming). TIDAL is available in fifty-three countries, with over 60 million songs and 250,000 high-quality videos in its catalog. Along with music track, TIDAL makes available to its users also original video series, podcasts, playlists curated by industry experts, music journalists, and artists. The added value that characterizes the TIDAL platform is therefore exclusive content alongside with unique experiences. TIDAL presents itself as a revolutionary platform, with a better distribution of revenues among artists and a commitment to the fight against piracy. The peculiarity of the service is that it offers a better experience in terms of quality, in fact, the tracks have a better audio quality than the competition. It also offers a lot of exclusive content and various special features, such as *Track Edit* that allows you to change the time and length of songs, the ability to interact with artists, follow streaming concerts and distribute tickets (Leocata, 2017).

SoundCloud (e) is the world's largest audio and music platform, allowing people to discover and listen to a selection of music from the most diverse community of songwriters on the web today (SoundCloud, 2020). Since its launch in 2008 in Berlin, the platform has become renowned for its unique content and features, including the ability to share music and connect directly with artists, as well as discover innovative tracks, raw demos, podcasts, and more. This is all thanks to an open platform that directly connects songwriters and their fans around the world. It is a very popular platform for independent and emerging artists. In 2017, it boasted 175 million active users with 125 million tracks. Unlike other streaming platforms, SoundCloud can rather be defined as a community where artists have the freedom in publishing unofficial content and fans have the opportunity to follow and interact with them. Thus, SoundCloud has a nature that lies in between the concepts of music platforms and social media apps. It offers the ability to synchronize one's account with that of a social media such as Facebook, Twitter, and YouTube.

A typical feature of SoundCloud is that songs can be commented on at a precise second of their duration. The number of contents is much larger than the competition because it is independent and unofficial material. Anyone has, in fact, the possibility to share audio files, and it is an excellent showcase for emerging artists. By choice of the distributors, the possibility of downloading the audio file can also be included.

Like some of the above platforms, it offers the possibility of free or paid use, with the advantages in the second case of having additional services and more space for uploading audio files. In March 2016, SoundCloud Go, the platform's first streaming service, was launched, with prices and modalities in the average of the competition, thanks to agreements made with the majors (Leocata, 2017).

Deezer (f) is a music service of the media player Blogmusik SA that offers on-demand streaming of songs from numerous record companies and independent labels. It is considered among the main European competitors of Spotify; it was launched in France a couple of years before the Swedish platform (Longo, 2014). Today the service is offered in 183 countries. It offers unlimited access to a music repertoire of 56M tracks with high-quality sound on any medium (smartphone, computer, tablet, hi-fi, or car stereo) to more than 16M active users worldwide.
The platform was launched in 2006 initially under the name of Blogmusik, the platform did not have all the agreements to legally distribute music streaming and for this reason, it was closed in February 2007 by SACEM – in French *Société des auteurs, compositeurs et éditeurs de musique* – due to copyright infringement. It reopens under the new name Deezer.
Like many of the platforms mentioned above, Deezer offers a free and a premium paid option, with the standard figures and benefits of its rivals. It also offers a service for listening to music with better, uncompressed audio, called *Deezer Elite*, but it comes at a higher cost. It is integrated with Facebook and Twitter, and one of its special features is the function that lets you read the lyrics of the songs through a simple on-screen command (Leocata, 2017).

Amazon Prime Music (g) is an online store of music content launched for the first time in January 2008 in the United States managed by Amazon.com. Since the end of 2008, the platform began to spread across the European continent. Differently from the other competitors, Amazon Prime offers a complete service to the consumer providing at a fixed annual price, free access to e-books, movies, tv shows, and music repertoire (Longo, 2014). In January 2020, Amazon Music had 55 million listeners. The Amazon Prime package also provides the possibility to use Amazon's free shipping service for purchases on the platform – unlike other streaming platforms, Amazon does not specialize only in music, but the service is diversified.
In 2016 Amazon developed *Amazon Music Unlimited*, a digital music streaming service as an additional tier to Amazon Prime Music or as a standalone subscription; it gives users access to more than 50 million tracks, playlists, and radios created by Amazon Music experts (Amazon Music Unlimited, 2020). With Amazon Music Unlimited users can listen to millions of songs anywhere

and anytime on all their devices: smartphone, tablet, PC, all without advertising. It is also included in the service the offline mode and personalized suggestions based on the consumer's preferences. The difference between the two subscriptions is that with Amazon Music Unlimited, users have all the benefits of Prime Music, but with a larger catalog of over 50 million tracks instead of 2 million tracks, and hundreds of playlists and radios created by music experts. On September 17, 2019, Amazon Music announced the launch of Amazon Music HD, a new tier of lossless quality music with more than 50 million songs in *High Definition*, and millions of songs in *Ultra High Definition*.

Google Play Music (h) ***and YouTube Music*** (i).

Google Play Music was launched in 2011 in the United States, and then since November 2012 in Europe (since 2017 also in Italy). The peculiarity of the platform was the option to upload up to fifty thousand songs from the hard drive, and then convert their CDs into digital format. The strength of the platform lay in the popularity of the Google brand and the advantage of the cloud, which is widely used by users. On May 12, 2020, Google announced the closure of the platform and that starting from that date, Google Play Music users would be able to migrate their playlists and songs within the new YouTube Music service.

YouTube Music represents a new streaming music service, app, and computer product reinvented with official albums, playlists, and singles (YouTube Music, 2020). It was first developed as an extension of the YouTube audio-video platform, and now YouTube Music has become a separate subscription service, positioned as a more direct competitor to platforms such as Apple Music and Spotify, offering ad-free and background-only audio streaming and downloads for offline playback, for music content on YouTube. The YouTube Music subscription has a fixed monthly price tag in line with its competitors. As of March 2019, the app is available in 43 countries.

2.4. Business models and use of machine learning

As mentioned above, the difference between downloading and streaming lies in the fact that in the first case the downloaded file remains in the possession of the user while in the second one is possible only the online fruition. Therefore, streaming platforms do not sell a product to the consumer but rather a *service*; listening to music online anywhere and anytime with the only limitation of having an internet connection. This distinction defines streaming platforms as *access-based* platforms differently from those of download that are defined *ownership-based* platforms.

Different types of streaming platforms differ according to the models they adopt for the remuneration of the service they give to the user; there are therefore three subcategories of streaming:

- *Subscription-based*: this model allows users to listen to music through the payment of a fixed monthly or annual fee. This type of platform does not allow free streaming for users but is based solely and exclusively on subscription. Platforms can offer different packages with different price ranges according to the proposed features. The reference example is TIDAL platform, a subscription service that combines music tracks in lossless format and high-definition video.
- *Supported by advertising*: final consumers do not pay to listen to music, the streaming platforms are financed by companies that request advertising space on them. With the development of Pandora Internet Radio, the idea behind it was to create a different radio station for each user with the music they preferred without having to listen to the genres that others were interested in. This platform has decided to be consumer-oriented by allowing free listening: in fact, revenues are almost entirely due to commercials present every few songs played.
- *Two-tier freemium model*: this model offers consumers two different modes of use; one is the free-of-charge service and the second is the flat-rate service. The free service does not require any payment from the consumer but is supported by advertisements. The revenue for the platform and consequently the remuneration of artists and record companies are generated through commercial breaks between songs and are supported by the advertisements that the user is required to listen to, without the possibility of skipping – approximately about every thirty minutes. The second mode is a premium service, a premium subscription to the platform, that charges consumers a monthly flat fee with a fixed installment and allows users to take advantage of additional benefits compared to the basic free version. The premium service usually offers, in addition to the basic version, the ability to listen to unlimited music without commercial interruptions, to save and later listen to offline playlists, to have unlimited skips, to listen to songs by choosing the desired track, and applications for different devices, e.g., mobile devices (Thomes, 2013). Platforms like Spotify, Deezer, and SoundCloud are examples of this type of business model. The vision of this kind of platform aims at eliminating piracy by convincing consumers all over the world to use a legal service that recognizes the right reward to the production chain of the music world and especially to the artists and musicians. The royalties paid to the record industry are based on an innovative system of remuneration called *pay-per-stream*: a payment is granted for each stream recorded by a song (Gramolini, 2015).

In order to enrich the music listening experience and to make it more and more personalized for consumers, streaming platforms make use of *AI – IT tools* that allow them to offer a personalized and unique service for each user of the music platforms. Among these tools there are:

- *"Artificial intelligence (1) that creates algorithms enabling the creation of customized songs for users and helps artists to focus more on being creative.*
- *Machine learning (2) that enables consumers to draw on past information, leading to increased trust among stakeholders.*
- *Fintech (3) the rise of the blockchain and bitcoin creating new methods of sharing, creating, and selling music.*
- *Virtual reality (4) artists can create interactive virtual worlds, allowing fans from all over the world to share experiences and open up new worlds and also enabling disabled (financially and physically) people to enjoy live music.*
- *Big data analysis (5) provides sources for real-time personalization by compiling wide-ranging personal information (e.g., purchasing history, listening habits, physical and mental conditions).*
- *Social media (6) exploring new distribution channels (e.g., Facebook, Twitter, YouTube)"* (Naveed, Watanabe and Neittaanmäki, 2017).

Following are some examples of the use of these IT technologies used by streaming platforms. *The Echo Nest* purchased by Spotify is a Boston-based analytics startup that has improved the analysis of music-related data, this technology uses software to analyze and classify music based on different audio factors. *The Echo Nest* also performs the function of analyzing online conversations about music that take place every day, around the world on blogs, review sites, tweets, and social media comments to then turn them into usable data on a quantitative level. Another example of the application of AI and IT tools on Spotify is its flagship *Discover Weekly* service, a service that recommends to the user the music that fits their context, based on several factors such as the listener's location, to the content they are consuming, and their current emotional level (Prey, 2017). The service *Discover Weekly* creates a personalized and unique playlist of thirty songs delivered to each user every Monday morning. The playlist is created through the use of artificial intelligence and machine learning algorithms. Following thorough data collection, users' music preferences, streaming history, or how many times they've listened to a particular song are recorded, then the system creates a unique playlist for each user that aims at meeting listeners' interests and tastes. These technologies allow the Spotify platform to control whether or not a song is listened to entirely, or even skipped. The more users listen to music, the more data the platform acquires about them and the better its algorithm specializes. *Spotify Home Screen* uses machine learning algorithm

known as BaRT. *"*BaRT *is a* Bayesian Additive Regression Trees *which is a Bayesian "sum-of-trees" model where each tree is constrained by a regularization prior to being a weak learner, and fitting and inference are accomplished via an iterative Bayesian backfitting MCMC algorithm that generates samples from a posterior. In Spotify, BaRT is used to predict the wide range of different shelves and shelf could be made for you or recommendations related to recent listening history"* (Nigam, 2019). BaRT algorithm is optimized for more than thirty seconds streams meaning that a song will be considered relevant only if a user listens to it for more than thirty seconds. *"The platform will then retrain the model once a day based on interaction data collected, and finally it will build the system to de-bias for positional bias, meaning that clicks on the top are considered less worthy compared to the ones to the bottom".*

Another example is the algorithm developed by the Music Genome Project used by Pandora. It analyzes four hundred different parameters of the song entered by the user, which must be present and cataloged in the Pandora archive. The playlist that is processed can then be listened to as a radio, or managed by the user himself who, of each song, can decide whether to skip it or mark it as *liked* or *not liked*. This will help the system in the future in proposing music the user will appreciate. The algorithm selects songs according to music characteristics and not according to the popularity of the artist or its sales.

2.5. Artists' payout

There are several ways in which streaming platforms pay artists, musicians, or record companies who are the rights-holders of songs.

- *Mechanical Royalties* are royalties paid to songwriters or publishers in order to obtain the right to reproduce a particular song. The name derives historically from the fact that these royalties were used to be paid for the mechanical production of the composition, but in the present streaming world, they are made valid if a person simply chooses to play a song.
- *Public Performance Royalties* are royalties paid to songwriters and publishers for the right to publicly perform music. Considering that the music on streaming platforms cannot be considered owned by a user; each stream is considered a public performance.

Royalties for mechanical and public performances are subject to strictly local legislation and may differ from country to country.

- The largest and most substantial portion of the overall payout is due to *copyright owners* on the recording side through labels and distributors. For these types of royalties, almost all

streaming services calculate this last part of the payout using the same rules (Soundcharts, 2020).

It has been calculated – updated in January 2020 – that *Tidal* pays $0.01284 per stream, *Apple Music* pays $0.00783 per stream**,** *Deezer* pays $0.0064 per stream**,** *Spotify* pays $0.00437 per stream**,** *Amazon* pays $0.00402 per stream**,** *Pandora* pays $0.00133 per stream**,** and YouTube pays $0.0.00069 per view (Dittomusic, 2021).

CHAPTER 3. The phenomenon of piracy and peer-to-peer

3.1. The digital revolution and the piracy phenomenon

It is a known fact that the advent of the digital age has disrupted and revolutionized the world and the way in which many industries were conceived and operate, it has accelerated technological progress bringing with it innovations and advantages that are still carried on today and that still drive the process of revolution constantly in progress. Unfortunately, like in most of the great revolutions, there are two sides of the same coin and together with outstanding advantages, some problems have also emerged. Copyright infringement is an example of the new challenges brought by the digital context and along with it also the rapid spread of piracy in the creative industries.

The digital revolution has been one of the major factors responsible for the illegal and unlicensed distribution of audio and video products as content has become easier to copy and to share with little to no loss in quality (Larabi, Rosselli, and Fernandez-Maloigne, 2009). It has given users the possibility to store information digitally allowing content owners to copy, keep, and transmit large volumes of information as originally formatted (Sudler, 2013). The introduction of the *World Wide Web* – i.e., www a commercial Internet protocol – has been a clear demonstration of how global distribution became rapidly easy and low cost. For these reasons, the digital revolution has become a "perfect storm" for online piracy – a condition that if not managed appropriately could significantly damage market growth and industry sustainability. The music industry, for instance, has been one of the most impacted by online piracy.

It is important to underline that even if after the diffusion of P2P platforms and the beginning of an era where, due to the introduction of the internet, digital piracy started to spread and to become more and more threatening to the creative industry, the phenomenon of piracy itself is not new. Piracy was not born precisely with the digital revolution and the advent of the internet, but it was present for a long time and was already a threat to rights holders. Already with the shift from live to recorded music industry, there are hints of the phenomenon of piracy. In particular, cassettes and CDs, are an example of how popular this phenomenon was. In fact, they were duplicated as many times as desired, for clandestine sale or simply for distribution among friends, relatives, and acquaintances. The advent of the internet has simply increased disproportionately this already existing practice as it has connected or brought together an infinite number of users who were geographically distant. Anonymity itself has fueled consumer complicity, permitting users to search freely for content without fear of getting caught or retributed. Furthermore, technological shifts allowed global access to pirated goods and extended access across international borders. Piracy also

pointed out the impact that technology has had on consumer's purchase decisions, specifically perception, product availability, cost, and quality. They have all changed in large part due to the digital and Internet paradigm shifts.

The beginning of the piracy phenomenon related to file-sharing systems instead can be traced back to the birth and spread of the Napster system. Napster, as previously mentioned, was a file-sharing program created by Shawn Fanning and Sean Parker. It was active from June 1999 until July 2001 and despite its short period of life has had such popularity to change forever the dynamics of the music industry. Napster was officially launched in 1999, it was the *first* mass peer-to-peer system, and since then, nothing has been the same. The software allowed users to search for as many songs as they want, as long as they were on one of the computers using the program, they could download them for free and then share them with everyone else. The growth was exponential: already in October 1999, Napster had 4 million songs; in March 2000 – less than a year after its birth – there were more than 20 million users of the software. In the summer of 2000, fourteen thousand songs were downloaded every minute. As expected, the music industry reacted to the threat: the Record Industry Association of America *RIAA*, Metallica, Dr. Dre, and several other big names in the musical world sued Fanning and Parker directly in order to prevent the use of the program, contesting the accusations of copyright infringement. The courts decided that Napster had violated the rules of the *Digital Millennium Copyright Act* and issued an ultimatum: *"remove all copyrighted content or close the program"*. The two founders were unable to comply with the court's demands, and in July 2001, two years after its inception, Napster servers were ordered to shut down due to repeated copyright infringement. Napster was ordered to pay $26 million as compensation for past damages and $10 million for future royalties. Napster in just two years of operation, managed to reach about 20 million people, resulting in losses in terms of earnings for the record companies.

Despite its closure, piracy continued to spread due to new, more sophisticated P2P software that did not need to rely on central servers, which made the illegal downloading of music a daily practice for a growing percentage of Internet users. Examples are eMule, Kazaa, Gnutella, Emule, DC++, Audio Galaxy, and Soul Seek. On July 2, 2001, just when Napster was closing its doors, the 26 years old Bram Cohen launched on the Internet what is still today the most important file sharing tool in p2p mode: BitTorrent. This resulted in continuous declines in revenues, with a 50% contraction until 2014. At a social level, there was a growing conviction that there was no need to buy music because it could be easily retrieved for free on the internet, so consumption behaviors and attitudes completely changed.

There have also been attempts to combat piracy by internationally renowned artists e.g., the association *Artists Against Piracy*, composed of 70 artists, that has been created with the purpose of

raising awareness of the issue of piracy, trying to make people understand the value of music with various advertising campaigns in major American newspapers and emphasizing the lack of control of artists over their works, with repercussions on their careers (Leocata, 2017).

If the music industry thought it had solved its nascent problems by forcing Napster to close, it had miscalculated. The effects of the multiplication of P2P platforms immediately begin to be felt: the worst year was reached in 2014 when global revenues plummet to 14.3 billion. In 15 years, the spread of piracy has halved the revenues of the music business.

In those years the music industry led a battle against the use of P2P platforms, tracking down every filesharing site or program that re-emerged under a new name a few days after being forced to close. The same thing still happens today with the many *proxies* that allow access to the most important site for Torrent distribution: *The Pirate Bay*. The criminalization of piracy did not bring any results.

In 2010, however, something starts to change: In Sweden, the birthplace of The Pirate Bay, the number of people downloading illegal files dropped by 25% from 2009 to 2011. In Norway, the drop between 2008 and 2012 is even 80%, from 1.2 billion illegal file downloads to 210 million. Over the following years, many nations began to follow the same trend. This decline in the use of P2P sites was due to the introduction and spread of platforms such as Netflix and Spotify, which slowed the spread of illegal file-sharing offering a legal alternative to systems like Napster and its clones. A report commissioned by the Swiss government pointed out that efforts to combat piracy have cost more money than they have earned. In contrast, the decline in music revenues was halted thanks to the emergence of legal alternatives to Napster, streaming platforms.

3.2. File-sharing architectures

File-sharing is the process of collecting digitized archives of information, documents, or electronic pieces e.g., audio and video content, with the aim of creating a database that can be shared online within platforms defined as computer networks. It is defined by the action of sharing content between computers connected to the same network, it refers to the computer activity of sharing files within a computer network. File-sharing in a few words is a decentralized and delocalized distribution system. File-sharing sites or platforms are programs that allow files to be exchanged between Internet users via a common server. This interchange through the web can be achieved through a *client-server* (client and servant) or *peer-to-peer* mechanism.

Client-server and peer-to-peer

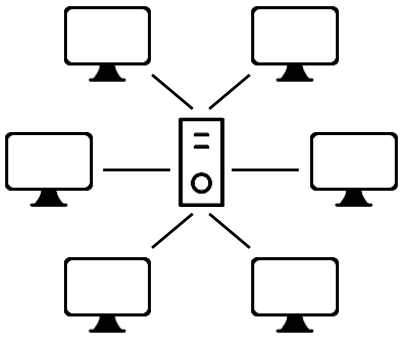
In a client-server network, computers called *servers* make available resources and offer services to other computers, called *clients*, while in peer-to-peer networks, computers play simultaneously both the role of client and server. The distinction in a client-server network is clear, i.e., the client cannot become a server and vice versa, it is called a *dedicated* server.

In the client-server model, communication occurs through the exchange of messages. A message is nothing more than a set of data that goes to constitute a complete entity. This means that users will be able to use the same software as they would on their own. When a connection is established between client and server, and the client requests the launch of an application, there are two possibilities: the execution can be client-side, i.e., the program is downloaded and executed on the personal computer, or it can be server-side, i.e., the software is executed directly on the server which then communicates the results to the client.

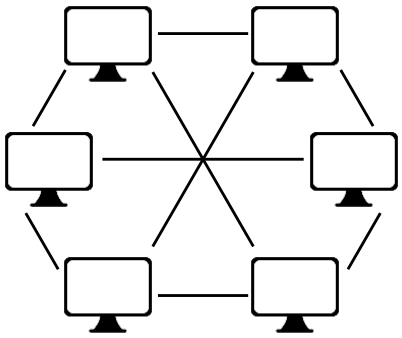
In the peer-to-peer model, there is no difference between client and server: the network is equal, and each computer makes up a node, i.e., both client and server.

Below, the difference between the two multichannel systems *client-server* and *peer-to-peer*.

Graph 7. Client-server and peer-to-peer architectures.



Client-server



Peer-to-peer

Peer-to-peer is the basis of file-sharing, it has the advantage that it is easy to manage but it has several disadvantages:

- There is no central control of network management and there is also no need to hire a dedicated network administrator.
- Each user is responsible for what he or she wants to share with others, which makes it difficult to control what is shared.

- There is no control over security; each computer uses its own data protection measures.
- The network becomes increasingly complex to manage, especially if the number of computers is high. In fact, peer-to-peer networks work best in environments with less than ten computers.

Putting on hold for a moment the technical explanations of how the different architectures used by file-sharing software work, let's return for a moment to focus on the impact that the spread of file-sharing has had on the creative industry, particularly on the music industry. File-sharing sites have led to the disintermediation of the music industry: in fact, through file-sharing, users exchange among themselves the various music files without turning to the industry players.
P2P programs, such as Napster, are interfaces that facilitate the sharing of any type of material, connecting easily and quickly to the database; they are perfect for sharing MP3 music tracks that are small in size and travel without any problem. In the nineties a lot of P2P software like BitTorrent, DC++, and SoulSeek started to spread, characterized by a further decentralization, that is the information collected about the users are not on one centralized database as in the case of Napster, but they are created and sent in real-time by the software that monitors the connected users, collecting information about the files they want to share on the net (Gramolini, 2015).
In this dynamic and changing environment, the majors found themselves joining together as victims of an unexpected phenomenon to which they did not know how to react. In a market where free downloads from the net were proliferating and the number of CD copies sold was decreasing, they proposed to stop this phenomenon by installing an anti-plagiarism software, *DRM (Digital Rights Management)*, able to protect original CDs from illegal copies or digitization of tracks. However, this was not a plausible solution in a cutting-edge technological field, which soon found a way to overcome this protection. Piracy does not only damage the revenue of the sector but also the jobs of the music industry. Statistical data reveals that the record industry, with the introduction of digital piracy, has suffered a loss in overall turnover of about 300 million euros.

3.3. Peer-to-peer sites' impact on the recording industry

Napster

As already mentioned in the previous chapters, Napster was created in 1999 by Shawn Fanning and Sean Parker, and with its introduction, it is possible to trace the increase in popularity and diffusion of the P2P file-sharing system. Napster was the first mass-market peer-to-peer system, it was active for only two years and ran from June 1999 until July 2001, but despite the short period

of activity, it had an incredible success that marked forever the trend of the music industry in the digital era. Based on the architectures described above, Napster's operation can be classified *"in the middle"* of the two systems; in fact, the program was based on a system of central servers that maintained the list of connected systems and shared files, while the actual transactions took place directly between the various users. The Napster software was shut down due to all the claims of copyright infringement, as it provided users with music content in a completely illegal manner without paying royalties to the copyright owners. In 2002, the brand and logo were acquired by the company Roxio, which re-launched the music download service as *Napster 2.0* as a paid legal streaming service – after a 30-day free trial. Today, *Rhapsody International Inc.* operates the Napster platform in thirty-three countries (Napster, 2020).

eMule

Following the closure of Napster, *eMule* was born on May 13th, 2002, created by the German programmer Hendrik Breitkreuz. eMule was developed as an open-source P2P file-sharing software, without centralized servers. Legally, eMule was stronger than Napster, indeed it survived until the beginning of 2007, unlike many servers at the time that were closed down. Its strengths were the simplicity and cleanliness of the graphical interface, the availability in several languages – about forty, and a large and active community of users. Many people still consider it one of the 'leaders' of illegal downloading in the 1990s, because of this eMule was one of the platforms that did the most damage to the music and film industry's revenue.

BitTorrent

After eMule, it was the turn of *BitTorrent*, which was born in 2002 as a protocol for the exchange of documents. It was studied by Bram Cohen and had a meteoric success. BitTorrent imposed a mechanism to coordinate the work of numerous computers, obtaining the maximum possible benefit for all; thanks to this system, each node contributed to the spread of the file. BitTorrent allowed users to distribute files of any type, the original document is fragmented into many small *pieces* that would be reassembled at the destination.

Between BitTorrent and the most popular peer-to-peer systems, there were two main differences: the first was that BitTorrent did not search for files by name: the user had to get a *.torrent* file from a dedicated website, moreover BitTorrent did not try at all to hide the last host responsible for the availability of a given file – as other platforms did. One of the significant disadvantages of BitTorrent compared to other peer-to-peer systems was that files died easily because it was a program designed more for spreading files rather than sharing them; this protocol offloads, through *"seeding",* the responsibility to other protocols. The method used by BitTorrent to distribute

files however resembled the one used by eDonkey and Kad – networks of *eMule* with some technical differences related to the connection of the various nodes and the duration of the files sent or shared.

The Pirate Bay trial

The Pirate Bay TPB is a file-sharing website based on the BitTorrent sharing protocol. It was created in 2003 in Sweden by Gottfrid Svartholm, Fredrik Neij, and Peter Sunde. According to the Los Angeles Times, in 2006 The Pirate Bay was *"the world's largest mediator of illegal downloads and the most visible member of a growing international anti-copyright or pro-piracy movement"* (Keller, 2006). Originally, The Pirate Bay allowed users to download BitTorrent files, small files that contain the metadata needed to download files from other users of different categories e.g., audio, video, apps, games, etc.
The Pirate Bay trial is a trial that took place in Sweden against four people accused of promoting copyright infringement through the torrent file search site The Pirate Bay. The prosecution was supported by a consortium of intellectual rights owners led by IFPI who filed claims against the owners of The Pirate Bay. Swedish prosecutors filed charges of « *administering, hosting and developing the site and thereby facilitating copyright law infringement to other persons* » on January 31, 2008, against Fredrik Neij, Gottfrid Svartholm, Peter Sunde, and Carl Lundström. Thirty-four cases of copyright infringement were listed, of which twenty-one involved music; the trial began on February 13th, 2009 and ended with the announcement of the verdict on April 17th finding the four owners guilty and sentencing them to serve one year in prison and pay SEK 31 million approximately €3,000,000. One of the biggest controversies of the whole episode was the consequences that The Pirate Bay trial brought afterward; namely, an increase of users to the pirate party following the trial, making it increase its popularity more and more, the refusal by Swedish ISPs – *Internet Service Providers* – to shut down the illegal site, and the *Bailout* operation organized by *Anonymous* against the IFPI international website. This was a *Denial-of-Service* attack on IFPI's site, making it unreachable for several hours, followed by a statement to P2P supporters that they would refrain from buying any media industry products in the months to come.

3.4. Creative industries copyright protection

Copyright protects intellectual works of a creative nature concerning science, literature, music, figurative arts, architecture, theater, cinematography, broadcasting, and, lastly, computer programs and databases, whatever the mode or form of expression (Treccani Italian Encyclopedia, 2020). Differently from industrial inventions, no filing fee is required for the protection of copyright.

In addition to copyright, there are *closely related* rights, which are intended to reward and incentivize the creative effort and investment of those who make such works accessible and usable by the public: performing artists, record producers, radio, and television broadcasters, etc.

The content of copyright is divided into moral and patrimonial copyright.

Moral copyright is a personal, inalienable, and intransmissible right. It is composed of a series of faculties, among which the right to claim the paternity of the work and to oppose any deformation, mutilation, or other modification of the work itself which could be detrimental to its honor or reputation. This right is inalienable and after the death of the author can be asserted, without time limit, by the direct ascendants and descendants (spouse, children, parents, brothers, and sisters).

Patrimonial Copyright consists of the exclusive right of economic exploitation of the protected work. It is composed of a series of powers, including the right to reproduce, distribute, communicate to the public, translate into another language, or rework the work. These faculties belong to the author or his assignees and have a limited duration in time since the exclusive exploitation can be exercised only for the whole life of the author and until the end of the seventieth calendar year after his death. All these faculties can be transferred, even separately, through a license contract or an assignment.

International copyright protection finds its main source in international treaties, the most important of which is the *Berne Convention for the Protection of Copyright in Literary and Artistic Works*, signed on September 9, 1886, and subsequently revised in Berlin, Rome, Brussels, Stockholm, and Paris (1971), it recognizes the original exclusive right of diffusion and exploitation.

However, it should be remembered that copyright is governed by the principle of territoriality, according to which each country has a distinct system of rules on this right; rules which have been progressively harmonized by international conventions since the end of the 19th century and by a large number of European directives since the early 1990s. The Ministry responsible for copyright is the Ministry for Cultural Assets and Activities - Directorate General for Libraries and Cultural Institutes - Copyright Office.

Copyright is an example of the new challenges posed by the digital revolution; as a result of increased user access to digital content and more opportunities to exploit copyrighted works, digitization has allowed for the spread of new forms of copyright infringement.

In the transition from the analogic to the digital world, copyright is probably the area that has undergone the most involution. On one hand, copyright has been hit by the criticism of those who see the exclusive right as a reward for innovation, since copyright is capable of spontaneous self-fulfillment, on the other hand, the numerous legislative measures adopted in recent years to prevent

online piracy have sometimes been perceived as forms of control and censorship of the free flow of information, as measures capable of threatening freedom of expression on the web.
Unfortunately, today the problem of copyright infringement, particularly in the digital world, continues. For this reason, several national and international bodies are promoting through awareness campaigns the importance of maintaining a healthy environment that benefits all parties, from artists to end users, from contributors to listeners, from labels to distributors, etc.

In the *Global Music Report 2020*, the IFPI body has outlined the fundamental principles that underpin the creation of a fair environment for the music industry. The report defines four pillars intending to specify what the behaviors and values of all participants in the music supply chain should be in order to ensure a sustainable environment and benefit all parties. These principles are called *"Four pillars of fair marketplace for music"* and are specified below.

1. **Music's value should be recognized** – *"Policymakers should recognize that music has both cultural and economic value. Rules should ensure that all services engaging in distributing music online, regardless of how they operate, negotiate licenses with right holders, those who create and own the music, in a fair, competitive marketplace"* (Moore, 2020).
2. **Copyright frameworks should be clear and provide** – *"A balanced and clear legal framework is needed to allow everyone to understand how music can be used legally. This should give right holders an adequate level of protection through exclusive rights, while allowing, in appropriate cases, clearly defined and targeted exceptions to those rights. Open-ended or 'flexible' exceptions are open to abuse and undermine this balance"* (Moore, 2020).
3. **All parties should be free to agree the terms of their relationship** – *"In a fair and functioning marketplace, parties should be free to agree the terms of their relationship. Unfair restrictions, whether over rights or contracts, distort and limit the development of music markets and result in recorded music being devalued"* (Moore, 2020).
4. **Adequate tools should be available to prevent music from being made available illegally** – *"As the online marketplace around the world continues to evolve, so too do the challenges the music community faces in preventing music from being made available illegally. There should be fair and effective ways to tackle illegal services that seek to exploit the work of artists and profit through large-scale copyright infringement"* (Moore, 2020)[2].

[2] IFPI, "*Global Music Report 2020, The Industry in 2019*".

CHAPTER 4. Literature Review

4.1. Literature review

The analysis implemented in this dissertation stems from an extensive study of previous literature covering the topics of streaming as digital technology, music streaming and music streaming platforms, copyright regulation and copyright infringement in the creative industries, piracy and illegal file-sharing platforms, and logistic regression analysis applied to these topics.
Following an extensive review of the literature, an attempt is made to analyze if there is a gap that could be filled and then an analysis is devised that could fill these possible gaps and make its own contribution to the total amount of research done on the topic. The focus of this research aims to establish the relationship between the introduction of music in digital formats and digital piracy. Additionally, the thesis research whether engagement with a legal streaming platform leads to an increase in industry revenue. It attempts to fill this gap by analyzing the factors that influence a consumer's WTP of a music streaming platform by identifying several variables that influence it.

Several studies have shown that countries with higher levels of corruption and poverty in the economy also show higher piracy rates on average. In 2016, MUSO conducted a study in order to monitor the presence of piracy around the world and track the habits of pirates. However, it turned out in the study that Europe has the highest rate of piracy; specifically, the nations that use it the most are Latvia (46.33%), Bulgaria (27.43%), and Lithuania (24.54%) (Grasso, 2016). Marron and Steel (2000) and Silva and Ramello (2000) showed that per capita income is negatively related to piracy rates. Chiang and Assane (2009) demonstrated that income, in particular disposal income, has a significant influence on consumers' willingness to pay for digital music. Chen, Chen, and Yeh (2010) investigated the determinants of software piracy in the Far East during the period 1995 to 2006. Their results show a negative relationship between software piracy and the unemployment rate, per capita GDP. A study conducted by Kyper, Prante, and Schimmoeller (2016) shows the role of income in predicting national software piracy rates. Andrés (2006) investigated the extent to which income inequality influences national piracy rates across a sample of 34 countries. The study resulted in a negative significant effect of economic inequality on national rates of piracy. Moreover, results show that judicial efficiency affects piracy rates. Bekir (2017) analyzed the piracy-corruption relationship, distinguishing between the direct impact of corruption on piracy and the indirect impact that operates through the impact of corruption on per capita income. Results show a positive direct effect of corruption on piracy rates.

Building on insights from previous studies, the following hypothesis about the role of discretionary income on WTP for digital music is posited:

H1: A higher degree of *discretionary income* is likely to increase *WTP* for digital music.

Cheng, Sims, and Teegen (1997) and Cohen and Cornwell (1989) researched, in their studies, whether piracy was related in any way to demographic factors, risk preferences, peer effects, income, and knowledge of copyright laws. Results have shown that as solutions against piracy there are law enforcement e.g., penalty applications, and economic incentives (Conner and Rumelt, 1991; Varian, 2005). Regarding enforcement, some studies view it as a measure to reduce piracy, while others do not. Yoon (2002) showed that enforcement leads to a private cost to consumers that discourages them from acquiring music illegally. Stolpe (2000) argued it can only be an effective measure if businesses can economically afford to adopt these guidelines. In contrast, Ben-Shahar and Jacob (2004) argued that selective copyright reduction increases a firm's market share in legal sales. Shy and Thisse (1999), Gayer and Shy (2003), and Peitz (2004), argue that copyright enforcement reduces network externalities. Marron and Steel (2000) and Gopal and Sanders (1998) showed that countries with stronger legal systems exhibited lower piracy rates. Becker (1968) studied the influence of risk factors, internalizing risk perception in economic functions to study whether they influenced engagement in piracy. Risk perception has been shown to be influenced by peer effects. Chiang and Assane (2009) in their empirical analysis on estimating willingness to pay for digital music included risk perception as a function of the peer effect among the variables. Several studies show how (illusorily) less risk is perceived if groups of individuals participate together in an event – illegal/dangerous practice. Danaher, Smith, and Telang (2013) concluded in their study that it is difficult to determine the socially optimal set of government copyright strategies and policies in the digital age. A study conducted by Suduc, Bizoi, and Filip (2009) demonstrates as security policies and codes of ethics combined with users' education/training measures are useful in deterring piracy misuse. Baird (2018) examined the willingness to pay for a music service among college students after being exposed to a digital advertisement. More specifically, Baird studied how rational and emotional appeals in advertisements influence interest and willingness to purchase a paid music service. Wang, Zhang, and Ouyang (2005) outlined a conceptual framework for understanding consumer ethical decision-making associated with piracy rates. They examined the effects of two culture-related constructs: assumption of responsibility and attitude toward copyright laws on consumer ethical decision making. A study conducted by Alleyne, Soleyn, and Harris (2015) shows a very favorable attitude, percentage-wise, towards illegal music; concluding

that piracy is easy and ethically accepted among students (Alleyne, Soleyn, and Harris, 2015). Lysonski and Durvasula (2008) suggested that participants in P2P music acquisition activities are unaware of the effect of illegal downloads on artists, musicians, and record labels. Moreover, their results show that illegal downloading continues at a high rate, driven by a strong belief that it is not ethically wrong. Ethical orientation was found to be positively associated with awareness of the social cost of downloading, consequences of downloading, and ethical belief in downloading. In addition, fear of consequences appears to impact the propensity to download illegally. Based on previous studies, this analysis assumes that:

H2: A higher degree of *ethics* is likely to increase *WTP* for digital music.

H3: A higher degree of *risk aversion* is likely to increase *WTP* for digital music.

Some researchers argue that the acquisition of music through illegal channels, P2P platforms, such as Napster, do not pose a danger to the sustainability of the industry. Alexander (2002) argues that there is not enough evidence to prove that P2P platforms have had a negative impact on the music industry's revenues (Alexander, 2002). Also, Aguiar and Bertin (2016) argue that music piracy does not have a negative impact on digital music purchasing behavior, and even though it involves the infringement of copyright it does not affect digital music revenues (Aguiar and Bertin, 2016). They contend that consuming illegal music could prove beneficial, in that it could bring consumers to buy it in a legal way (Aguiar and Bertin, 2016). The extent to which WTP for copyrighted goods increases due to piracy as a result of network externalities has been analyzed also by Shy (2001) and King and Lampe (2003). It turns out that in some cases, as previously mentioned, piracy can positively influence the willingness to pay. A study conducted by Pham, Dang, and Nguyen (2020) aimed to discover the factors that influence digital piracy behavior in Vietnam. The results showed that perceived behavioral control is influenced by technological development and perceived risk. A survey concerning the history of digital piracy is implemented by Belleflamme and Peitz (2010). The work of Thomes (2013) provides a theoretical analysis of a business model that offers the required music as a stream to potential customers. Using a free-entry model, Choi (2006) tested the inverse effects of the number of new market entrants and the level of advertising on social welfare.

A study by Sinclair and Green (2015) showed that music streaming services are popular among participants aged 19-30. Cheng, Sims, and Teegen (1997), Gopal and Sanders (1998), and Chiang and Assane (2007), used survey data on college students for surveys related to music consumption. Following the same logic, for this dissertation as well, the survey was distributed to a large

percentage of college students as they represent a market segment that is akin to technology and influenced by the peer effect (and in particular, consumers aged 18-35 are considered). Chiang and Assane (2009) use peer effect variables to study the influence of users' environment. Specifically, they focused on a user's main area of study. Similarly, in this analysis the variable of *mastery of hardware use* is used to investigate users who are more in contact with technology or technology-inclined; the intuition behind this variable is that these users may find new technologies for acquiring music easier to pursue, which could lead to a lower WTP for purchased music.

H4: A higher degree of *mastery of hardware use* is likely to decrease *WTP* for digital music.

Kwong, Yau, and Lee (2003) observed that more than seventy percent of their sample of respondents had purchased pirated CDs in a year. Gupta, Gould, and Pola (2004) found that more than forty percent of their study participants admitted to software piracy. In a study by Lysonski and Durvasula (2008), almost ninety-five percent of participants admitted to consuming downloaded music for which they did not pay. Veitch and Constantiou (2011) highlight how the decision between legal alternatives and piracy has received limited attention in the literature. Therefore, in their article, they address these issues and attempt to fill these gaps by presenting a model of a user's decision to purchase digital products in the context of piracy. They introduce the concept of *"acquisition decision"* which is influenced by the user's product desire, price perceptions, perceived risks, internal regulators of behavior, resources, and product availability (Veitch and Constantiou, 2011). As reported, the literature regarding the relationship between legal music consumption and pirate downloading is not substantial, for this reason, the study carried out in this dissertation is interested in analyzing the relationship between music streaming-digital platforms and other channels of consumption; assuming that:

H5: A higher degree of *consumption through other channels* is likely to decrease *WTP* for digital music.

Armstrong and Weeds (2007) show that the quality of a platform is higher with paid subscription access than with advertising funding. Articles by Anderson and Coate (2005) and Peitz and Valleti (2008) investigate the social optimization of advertising versus the annoyance of advertising. To conclude, the variable *convenience* has been inserted in this analysis; it aims to establish if the convenience in using streaming platforms – i.e., if the ease of use and/or the features that a platform offers – influence a user's willingness to pay for the platform itself. The hypothesis underpinning the relationship that is researched in this study is:

H6: A higher degree of *convenience* is likely to increase *WTP* for digital music.

As it can be perceived from this brief summary of previous studies and research, there is still no clear relationship between the introduction of music in digital formats and digital piracy. Besides, it has not yet been fully demonstrated to what extent engagement with a legal streaming platform leads to an increase in industry revenues, resulting in an improvement in the industry. This dissertation aims to fit these pieces together to understand the influence of various factors on each other, specifically the factors that influence the WTP of a consumer of a music streaming platform. This dissertation, therefore, analyzes variables designed to test their possible influence on consumer WTP for legal subscriptions to music streaming platforms. Some of the factors analyzed were taken from previous studies and research and repurposed for the specific case discussed here. Specifically, several variables such as risk and ethics have been taken and readjusted from a study by Chiang and Assane (2009). Since the two researchers' study was completed in 2009, when streaming platforms such as Spotify (2008) had just emerged, one limitation may be that not all effects were captured. This thesis sought to expand the study on the factors influencing the WTP for digital music, updating it to today's times and evaluating other variables and factors.

CHAPTER 5. Data, Methodology, and Model

5.1 Data source and collection

Data collection for this study was implemented through a quantitative survey conducted through a survey platform called *Qualtrics*. The survey was created in three different languages – Italian, English, and Spanish – and distributed via a direct hyperlink to colleagues, friends, acquaintances, and fellow students within a set age range of 18-35 years old. The choice to consider a restricted sample in terms of age of respondents was determined by the fact that the intent of the analysis was to consider a target of people with similar habits – concerning this topic. Underage respondents were excluded as they are not legally able to pay for an online membership, while older respondents were not considered due to the hypothesis of different approaches to the digital world; it was estimated that the age range 18-35 could yield consistent results across interviewees.

The survey consisted of 47 multiple choice questions with 3 initial screening questions. Respondents were allowed to complete the survey from any device that had an Internet connection. The survey was administered between October and November 2020. In total, 859 responses were collected (not considering partial responses that were eliminated), of which only 640 were considered in this analysis that correspond to responses from people who passed the initial screening stage. The survey took around 10 minutes to be completed and needed to be fully completed in order for the answers to be used. The answers have been kept anonymous and confidential, and the data collected have been analyzed in an aggregate form.

In the sample, 55% of respondents are women; the average age is 25 years old; 41% of respondents graduated with a bachelor's degree, and 27% with a master's degree. In total, responses were collected from 34 different countries; 83% of respondents were of Italian nationality.

Figure 1. Percentage of workers and students.

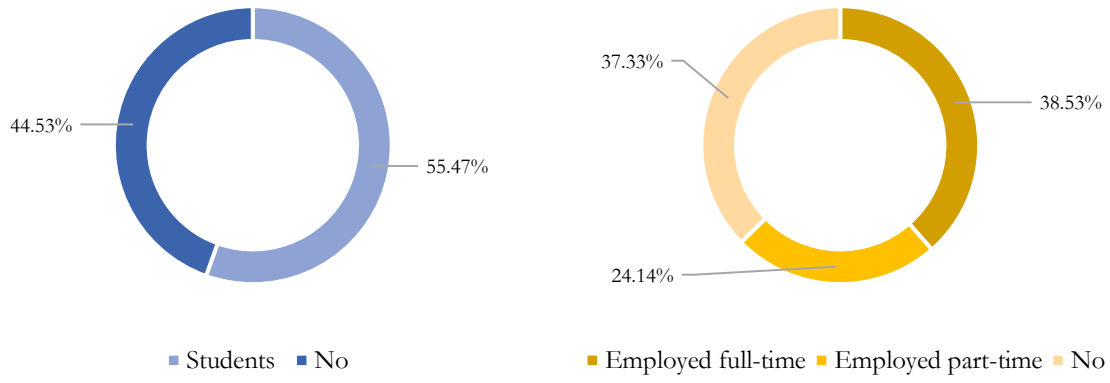
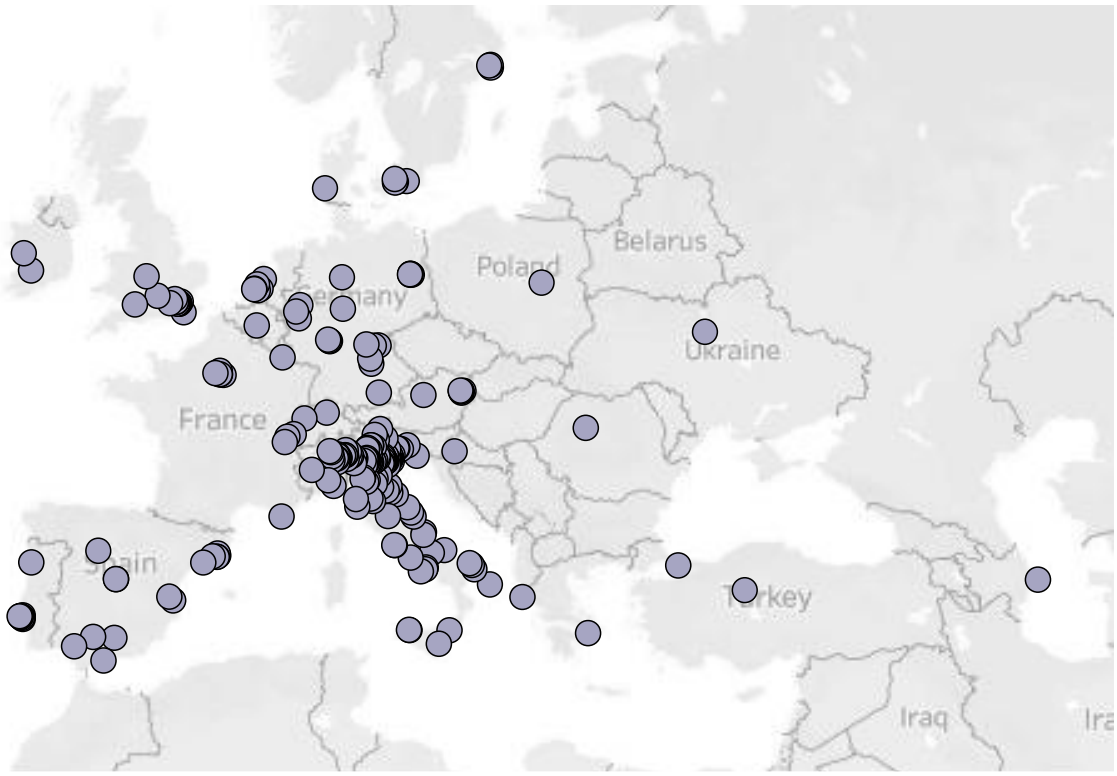


Figure 2. Responses collected by location (1).

Figure 3. Responses collected by location (2).



Screening questions are questions that allow a survey to be directed to a specific audience, filtering the respondents (qualifying or disqualifying them depending on their answers). In particular, this survey, in order to allow for a data collection that was consistent with the subject matter, targeted people who belonged to *all* three of these categories:

- people who listen to music with a medium to high frequency, for this study values considered were those equal to 3, 4 or 5, on a constructed scale of 1 to 5, with 1 for "never" and 5 for "always";
- people possessing a minimal knowledge of the concepts of piracy and P2P file-sharing sites;
- people with an understanding that downloading or uploading music to the platforms cited above is criminally punishable.

Respondents who were not on target were thanked for their time and the survey closed immediately after the three screening questions. The total number of people excluded from the survey through the screening question is 199.

The questions in the survey covered the topics of streaming, legal streaming platforms, and illegal P2P file-sharing platforms, piracy, music supports, digital, liquid, and live music and assessed people's knowledge and perception about these topics. The questions related to the variables have been structured by engaging a 5-point Likert scale applied to four different statements for each variable. Other questions have been inserted for statistical purposes such as demographic characteristics; nationality, age, gender. With the intent of obtaining truthful answers, the questions have been designed to avoid the use of terms with negative connotations – i.e., no direct questions related to the use of illegal platforms or sites were included.

5.2. Variables and Controls

Dependent Variable

Willingness to Pay for digital music

In this analysis consumers' behavioral intentions towards legal streaming platforms are examined through the study of their WTP for digital music. In economics, WTP stands for Willingness to Pay and indicates the maximum price at which an individual is willing to purchase a unit of a product. Normally WTP is defined as a precise value expressed by a consumer regarding its evaluations of a product, some researchers however consider WTP as a range of numbers.

In this study, since the dependent variable is treated as binary, thus accepting the two responses "*yes*" and "*no*", an individual's WTP is adjusted for his or her willingness to *subscribe* or *not* to the

service i.e., purchase (or not) a membership to a platform and pay (or not pay) for its subscription to the service. The WTP is adapted to the concept of WTP willingness to buy and is detached from its definition of expressing a maximum value to be attributed to a platform subscription and simply expresses in a more simplified way an individual's willingness to subscribe or not subscribe. The dependent variable in the model WTP (coded 0=no, 1=yes), defines the "do not subscribe" group as the reference baseline category, and the "do subscribe" group as the target category.

Independent Variables

The independent variables used in this analysis consist of economic and non-economic factors that influence the WTP for a subscription to digital music streaming platforms.

Discretionary income

The economic factor is measured by *discretionary income*, which is defined as the amount of an individual's income that is left for spending, investing, or saving after paying taxes and paying for personal necessities, such as food, shelter, and clothing; it includes money spent on luxury items, vacations, and nonessential goods and services (Investopedia, 2021). In the reviewed literature, several studies analyze the impact of direct income or disposal income; in this analysis, in order to capture the portion of income used exclusively for goods or services outside the necessity sphere, it was chosen to use discretionary income as an independent variable possibly influencing WTP. *Discretionary income* is used with the intent to provide insight into the ability or WTP for legal – copyright goods, with the presumption that a higher level of discretionary income is associated with less piracy or a lower tendency to consume or acquire content on P2P sites. In this model, the values for discretionary income were divided into five ranges expressed in euros. Respondents were asked to select the range of discretionary income that characterized them; the five ranges were expressed as values from 0 to 4, respectively 200 euros or less (0), 200 to 500 euros (1), 500 to 800 euros (2), 800 to 1000 euros (3), and above 1000 euros(4).

Ethics

The variable *ethics* is a variable composed of two sub-variables that are *fairness* and *shut* – described below – and it indicates the ethical component that pushes or not an individual to the consumption of products or services offered in a legal manner i.e., the payment for the streaming service. Consequently, the *ethics* variable is assumed to be indirectly proportional to the tendency to piracy and consumption or acquisition of music content on illegal P2P file-sharing sites. In this model,

the variable *ethics* is first considered by keeping the two variables fairness and shut separated, then it is also analyzed as the sum of the two. The hypothesis underlying the study is that higher values in the *ethics* variable (meaning higher values of *fairness* and *shut*) imply higher WTP for digital music.

Fairness

The variable *fairness* measures whether an individual believes piracy is unfair to copyright owners and damaging the music industry and its related parties. In this model, the *fairness* variable is analyzed as a continuous variable, and data were collected using the Likert scale technique on a range from 1 to 5 indicating the extent to which an individual agrees or disagrees with four statements regarding fairness to copyright owners, musicians, artists, etc. Following the values 1-5 with their respective meaning used in the survey; 1 for "Strongly disagree", 2 for "Disagree", 3 for "Neither agree nor disagree", 4 for "Agree", and 5 for "Strongly agree". After collecting respondents' preferences for the four statements proposed, the variable *fairness* is expressed in the model as the sum of the four values – i.e., fairness statement 1 + fairness statement 2 + fairness statement 3 + fairness statement 4.

Shut

The variable *shut* indicates whether P2P file-sharing sites and in general websites facilitating copyright piracy should be shut down because unsafe, unethical, and illegal. Analogously, the variable *shut* is expressed as the sum of the four statements proposed in the survey and each statement was measured individually with the same technique and the same scale used for the previous variable.

Risk aversion

Risk aversion is defined as the preference of an economic agent for a certain amount more than for a random one, i.e., the agent prefers to receive the expected value of a random variable over the value that the variable can assume. The valuation of the premium, however, is subjective and therefore is measured in terms of the *expected utility* that the premium confers on the individual; a risk-averse individual attributes greater utility to the expected value of a certain event obtained with certainty than to the expected utility obtained by participating in an uncertain event e.g., a lottery (Treccani Italian Dictionary, 2021). In this analysis, it is hypothesized that individuals who are more risk-averse by nature or those who are more likely to be influenced by risk warnings are more likely to reduce their participation in the illegal market i.e., participate in the file-sharing activity - piracy.

Similarly, to the *ethics* variable, *risk aversion* is also expressed as a sum of two sub-variables which are variables *caught* and *penalty*.

Caught

The variable *caught* corresponds to the perceived likelihood of an individual browsing illegal sites and acquiring music via P2P platforms to be caught and criminally prosecuted. The variable *caught* is measured similarly to those described above, on a scale of 1 to 5, with the Likert scale technique indicating respectively: 1 for "Extremely unlikely", 2 for "Unlikely", 3 for "Neutral", 4 for "Likely", and 5 for "Extremely likely". The values of the variable *caught* are expressed as a sum of the values of the four statements taken individually i.e., caught statement 1 + caught statement 2 + caught statement 3 + caught statement 4.
It is hypothesized that higher values of this variable indicate a tendency to move away from piracy and thus higher values of WTP – a positive coefficient is therefore expected.

Penalty

The variable *penalty* measures the perception of all the consequences that will befall a person if he or she is found guilty of a piracy offense e.g., illegal music download; it may include punishments such as fines, or prison sentences (even though they are rarely used in practice) and consequences such as bad reputation. As with the other variables, *penalty* is also measured using four statements and a 1-5 Likert scale indicating the values respectively: 1 for "Not at all concerned", 2 for "Slightly concerned", 3 for "Somewhat concerned", 4 for "Moderately concerned", 5 for "Extremely concerned". A higher level of penalty perception is expected to imply a higher level of WTP for a streaming platform subscription.

Mastery of hardware use

The *mastery of hardware use* variable – for simplicity referred to only as *mastery* – indicates the extent to which a person feels proficient and confident in the use of technology (particularly new technology). The intuition behind this reasoning indicates an inverse relationship between the *mastery* variable and the dependent variable WTP; this because it is expected that people who are in close contact with technology and feel proficient in its use are technically more likely to find new ways to acquire music more easily than "simply" paying for a subscription to a streaming platform. It is therefore expected that a higher value in *mastery* is matched by a lower value in WTP.

The *mastery* variable was also measured as a continuous independent variable with values on a Likert scale of 1-5 collected on four separate statements. The final value used in the model, as for the other variables, is the sum of the four values indicated in the four survey statements. Respectively, it is indicated 1 for "Strongly disagree", 2 for "Disagree", 3 for "Neither agree nor disagree", 4 for "Agree", 5 for "Strongly agree".

The variables *ethics*, *fairness*, *shut*, *risk aversion**, *caught*, *penalty, and mastery**[3] refer to a previous analysis conducted in 2009 in the United States by Eric Chiang and Djeto Assane, entitled *"Estimating the willingness to pay for digital music"*. The variables have been taken from the previous literature and readapted to the model of this analysis (Chiang and Assane, 2009).

Consumption through other channels

The variable *consumption through other channels* – or just *consumption* – analyzes the frequency with which a person gets in touch with music – i.e., listens to or acquires – through other channels other than streaming platforms. In particular, in this study, it is analyzed as "other channels" the radio listening (1), the purchase of tracks on online stores such as iTunes (2), the purchase of physical music such as CDs, vinyl, etc. (3), and the participation in events and live concerts (4). *Consumption* is again treated as a continuous variable, frequency is measured on a Likert scale of 1-5 and the values indicate 1 for "Never", 2 for "Rarely", 3 for "Sometimes", 4 for "Often", 5 for "Always". This study expects a negative relationship between WTP and *consumption*, in that an individual who makes frequent use of other channels for music listening may be less likely to pay for a subscription at an online streaming platform.

Convenience

The *convenience* variable measures the value that a person gives to the features that characterize streaming platforms and that simplify the use and listening to music. Features that make it more convenient to buy a subscription to a platform than to use illegal P2P file-sharing sites. These features include the ease – or intuitiveness – of using an online streaming platform, the ability to have different accounts and devices synchronized on the same library, the ability to connect social media apps and transmit tracks easily from one platform to another, the ability to listen to music offline in any device, anywhere and anytime, and the feature of having personalized suggestions of playlists based on the listening of a user and created ad hoc every week for the premium subscriber. It is therefore expected that users who positively and highly value these aspects – features – of a

[3] Variable* names have been changed and readapted.

streaming platform are more willing to pay for it. *Convenience* is measured exactly like all other independent variables – i.e., as a continuous variable – and 1-5 Likert scale values indicate respectively 1 for "Strongly disagree", 2 for "Disagree", 3 for "Neither agree nor disagree", 4 for "Agree", 5 for "Strongly agree".

The following tables, tables 1 and 2, summarize the variables described above indicating for each one the observations collected by the survey, the mean, the standard deviation, and their description regarding the values expressed in the model. The two variables *ethics* and *risk aversion* that are the sum of two sub-variables are inserted in the table together with their components but in the model, they have been kept divided. Therefore, the number of variables considered in the final analysis is six (of which two are composed of two sub-variables each).

Table 1. Variables, observations, mean, and sd.

	(1)	(2)	(3)	(4)	(5)
Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
wtp	**640**	**0.60**	**0.49**	**0**	**1**
ethics	640	28.15	5.50	8	40
fairness	640	14.50	2.83	4	20
shut	640	13.64	3.25	4	20
risk aversion	640	16.35	5.47	8	39
caught	640	9.42	3.15	4	20
penalty	640	6.93	3.91	4	20
convenience	640	16.18	2.54	6	20
mastery (hu)	640	15.59	2.97	5	20
consumption (oc)	640	8.71	2.17	4	18
discretionary income	640	1.35	1.48	0	4

Table 2. Variable, description, type.

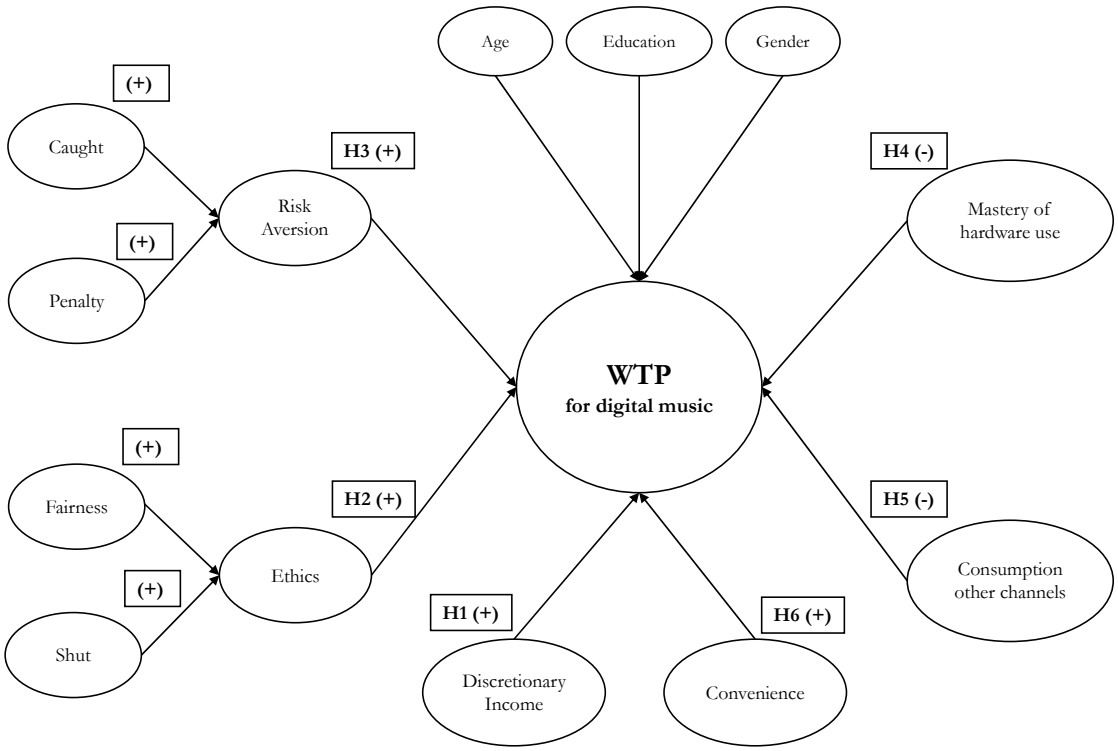
VARIABLE	(1) description	(2) type
wtp	**Willingness to pay for digital music – subscription to a music streaming platform. 0 for "no" 1 for "yes"**	**binary**
ethics	Sum of fairness and shut	continuous
fairness	1 for "Strongly disagree", 2 for "Disagree", 3 for "Neither agree nor disagree", 4 for "Agree", 5 for "Strongly agree". Sum of fairness 1 + fairness 2 + fairness 3 + fairness 4	continuous
shut	1 for "Strongly disagree", 2 for "Disagree", 3 for "Neither agree nor disagree", 4 for "Agree", 5 for "Strongly agree". Sum of shut 1 + shut 2 + shut 3 + shut 4	continuous
risk aversion	Sum of caught and penalty	continuous
caught	1 for "Extremely unlikely" 2 for "Unlikely" 3 for "Neutral" 4 for "Likely" 5 for "Extremely likely". Sum of caught 1 + caught 2 + caught 3 + caught 4	continuous
penalty	1 for "Not at all concerned" 2 for "Slightly concerned" 3 for "Somewhat concerned" 4 for "Moderately concerned" 5 for "Extremely concerned". Sum of penalty 1 + penalty 2 + penalty 3 + penalty 4	continuous
convenience	1 for "Strongly disagree", 2 for "Disagree", 3 for "Neither agree nor disagree", 4 for "Agree", 5 for "Strongly agree". Sum of convenience 1 + convenience 2 + convenience 3 + convenience 4	continuous
mastery of hardware use	1 for "Strongly disagree", 2 for "Disagree", 3 for "Neither agree nor disagree", 4 for "Agree", 5 for "Strongly agree". Sum of mastery 1 + mastery 2 + mastery 3 + mastery 4	continuous
consumption through other channels	1 for "Never" 2 for "Rarely" 3 for "Sometimes" 4 for "Often" 5 for "Always". Sum of consumption 1 + consumption 2 + consumption 3 + consumption 4	continuous
discretionary income	0 for "200 EUR or less" 1 for "200 EUR - 500 EUR" 3 for "500 EUR - 800 EUR" 4 for "800 EUR - 1000 EUR" 5 for "1000 EUR or more"	categorical

*

*Variables with response data that ranges from 1 to 5 (Likert scale) would also be considered ordinal variables.

As shown in Table 1, the mean and standard deviation of the two variables *ethics and risk aversion* are higher in terms of numerical value than those of the other variables because they are produced by the sum of two factors.

Graph 8. Conceptual model.



Graph 8 shows the relationships that have been estimated among the independent variables IVs (x) and the dependent variable WTP for digital music (y). At the top of the graph are included the three control factors that have been used in the model: *gender, education,* and *age.* The sign displayed next to each variable indicates the hypothesis that was made before running the model, e.g., if higher values of ethics are expected to increase the probability of high values in WTP the sign will be positive (+) and vice versa.

Table 3. Correlations of variables.

	wtp	*fairness*	*shut*	*caught*	*penalty*	conv.	mast.	cons.	discret.inc.
wtp	1								
fairness	0.27	1							
shut	0.19	0.64	1						
caught	-0.05	-0.01	0.09	1					
penalty	0.04	0.13	0.20	0.19	1				
conv.	0.36	0.30	0.25	-0.08	0.13	1			
mast.	0.20	0.12	0.03	-0.15	0.04	0.29	1		
cons.	-0.03	0.08	0.08	0.09	-0.06	0.02	0.04	1	
discret.inc.	0.10	0.09	0.01	0.02	0.01	0.00	0.09	0.05	1

	wtp	ethics	risk	conven.	mastery	consump.	discret.inc.
wtp	1						
ethics	0.25	1					
risk	0.00	0.16	1				
conven.	0.36	0.30	0.05	1			
mastery	0.20	0.08	-0.06	0.29	1		
consump.	-0.03	0.09	0.01	0.02	0.04	1	
discret.inc.	0.10	0.06	0.02	0.00	0.09	0.05	1

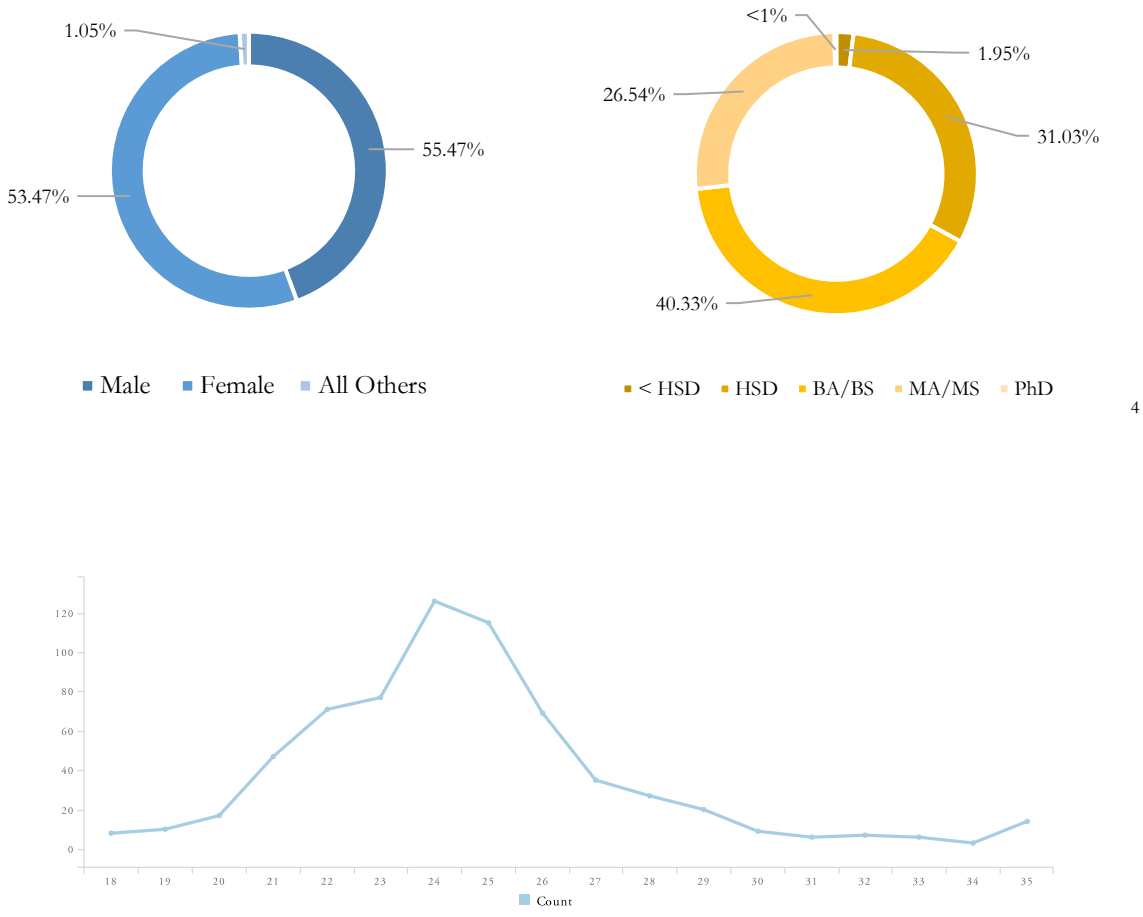
Once the variables for the model are defined, it is important to check that there are no cases of multicollinearity between them. Multicollinearity is a problem due to excessive correlation between two or more explanatory variables within a predictive model. This phenomenon can yield a model that is unreliable as the presence of multicollinearity means that two or more variables are explaining the same phenomenon leading to skewed or misleading. The correlation matrix in Table 3 shows the correlation between the variables considered in this model; as no values are close to 1, it can be stated that in this case, a problem of strong multicollinearity does not arise – indeed, weaker is the correlation, more reliable would be the model.

Control variables

Control variables are those variables that are held constant during an experiment, they are not part of an experiment itself (not independent or dependent variable), but they affect the results.

The three control variables – or simply just controls – that are included in this model are *age*, *gender*, and level of *education* achieved (called *education* for simplicity). The *age* control is a continuous nominal variable and the data collected are numerical values ranging from 18 to 35 (years old) – as for the sample, it was decided to consider people within this age range. As for the *gender* variable, it is treated as a dummy variable, taking the value 0 for "male" and 1 for "female" – a third value 2 was added to indicate the category "other" and "prefer not to disclose". The control *education* is divided into five levels from lowest to highest respectively: 0 for "Less than High School Diploma or Professional Certificate" 1 for "High School Diploma" 2 for "Bachelor's degree" 3 for "Master's degree" 4 for "Ph.D.".

Figure 4. Controls factors representation: gender (1), education (2), and age (3).



[4] <HSD for "Less than High School Diploma or Professional Certificate", HSD for "High School Diploma", BA/BS for "Bachelor's degree", MA/MS for "Master's degree".

Table 4. Controls, observations, mean, and sd.

	(1)	(2)	(3)	(4)	(5)
Control	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
age	640	24.63	3.09	18	35
gender	640	0.57	0.51	0	2
education	640	1.93	0.80	0	4

Table 5. Control, description, type.

	(1)	(2)
CONTROL	**description**	**type**
age	from 18 to 35 (in years)	continuous
gender	0 for "male" 1 for "female" (2 for "other – prefer not to disclose")	categorical
education	0 for "Less than High School Diploma or Professional Certificate" 1 for "High School Diploma" 2 for "Bachelor's degree" 3 for "Master's degree" 4 for "Ph.D."	categorical

5.3. Methodology

This study aims to test whether there is a relationship between several variables designed to represent different personality traits of music listeners and their willingness to pay for a membership to a streaming platform. This dissertation chose different types of variables to analyze that go into different spheres of the music listener e.g., personal values (*ethics*), tendencies (*risk aversion*), economic situation and spending preferences (*discretionary income*), consumption preferences, and utility (*convenience*), customs and habits (*mastery* and *consumption*), along with control factors as *gender*, *age* and level of *education* achieved. Analyses are carried out to identify when and if there is a relationship between consumers' willingness to pay and the IVs chosen for investigation. Once it is determined whether or not there is a relationship among the IVs and WTP for digital music – indeed whether each IV is relevant to the variable *y*, it is analyzed whether their relationship is, positive or negative, direct or inverse, to determine if it is consistent with the hypotheses

previously stated. With the particular aim of identifying if there are any factors influencing music customers' willingness to pay for the product or service they consume, the study also wants to shed light on the relationship between streaming platforms and piracy to understand, albeit in a simplified and summarized way, whether or not there is a connection between the two different music acquisitions. Given the limited presence of findings regarding this topic in the prior literature, this study tries to understand if these two different types of music acquisition have been and are substitutes or complements; i.e., if the introduction and spread of streaming music platforms have somehow been decreasing the illegal listening and downloading of creative content (Dang Nguyen, Dejean and Moreau, 2012).

5.3.1. Binary Logistic Regression Model

In order to analyze the data collected through the survey, this study makes use of a logistic regression model that is called binary logistic regression or binary logit model.
In statistics, the logit model, also known as the logistic model or logistic regression, is a nonlinear regression model used when the dependent variable is of a dichotomous type hence the name binary. It is used to model the probability of a certain event; the objective is to determine the probability with which an observation can generate one or the other value of the dependent variable. The logit model can also be used to classify observations, based on their characteristics, into two categories e.g., sick/healthy, alive/dead, pass/fail.
In a Binary Logistic Regression model, the dependent variable is dichotomous i.e., it can have as its only values 0 and 1 or values traceable to them; the regression calculates the probability that this variable acquires value 1.

Generally, in *logistic regression models* the relationship between a set of independent *Xi* can be:

- Dichotomous: e.g., yes/no
- Categorical: e.g., social class
- Continuous: e.g., age

and a dichotomous variable *y*.
In this study, the dependent variable *y* is represented by the dichotomous variable *WTP for digital music*, which takes the values of 0 and 1 for yes/no, the independent variables X_i are continuous (*ethics, risk aversion, convenience*, etc.)

In a logistic regression model, the dependent variable *y* defines the membership in one group or the other. Unlike in linear regression, what is of interest in a logit model is *not* the *predicted* or *expected* value, but the *probability* that a given subject belongs to less than one of the two groups.

In a binary response model, interest lies primarily in the *response probability*;

$$P(y = 1 \mid x) = P(y = 1 \mid x_1, x_2, x_3, \ldots, x_k)$$

where ***x*** is used to denote the full set of explanatory variables.
In the logit model, it is assumed that the response probability is linear in a set of parameters, β_j;

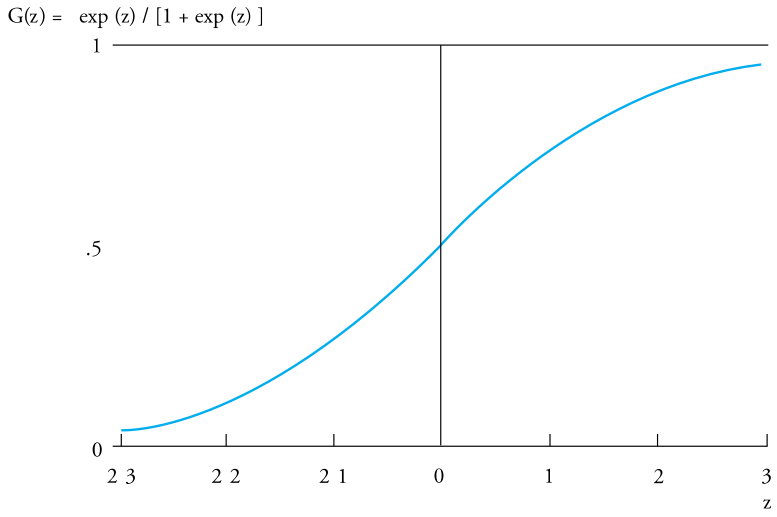
$$P(y = 1 \mid x) = G(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \ldots \beta_k x_k) = G(\beta_0 + x\,\beta)$$

where *G* is a function taking on values strictly between 0 and 1: $0 < G(z) < 1$, for all real numbers ***z***. This ensures that the estimated response probabilities are strictly between zero and one. In particular, in the *logit* model, ***G*** is the logistic function:

$$G(z) = \exp(z) / [1 + \exp(z)]$$

which is between zero and one for all real numbers ***z*** (Seddighi, 2013).

Graph 9. Logistic function $G(z)$ (Seddighi, 2013).



Logistic regression models estimate probabilities of events as functions of independent variables. Given one or more independent variables X_i called *predictors* (that as mentioned above can be either dichotomous, categorical, or ordinal), and a dichotomous dependent variable *y,* in the logistic model, the **log-odds** – *the logarithm of the odds* – is an estimation represented by a linear combination of the independent variables X_i – i.e., the log-odds for the value labeled "1" is a linear combination of one or more predictors. The corresponding probability of the value labeled "1" can vary between 0 (certainly the value "0") and 1 (certainly the value "1").

Given 2 mutually exclusive events, the log-odds is defined as the ratio of the probabilities that an event and the other will occur. Odds is a way of expressing probability by means of ratio; an odds ratio for a predictor is defined as the relative amount by which the odds of the outcome increase or decrease when the value of the predictor variable is increased by 1 unit. It is calculated as the ratio between observed frequencies in one event with the ones in the other. The odds ratio expresses the relationship between two categories as a function of another variable.

Ordinary Least Square OLS vs. Binary Logistic Regression

In an OLS regression, the relationship between the independent variables and the dependent variable is assumed to be linear, and the residuals are normally distributed and show a constant variable. However, when dealing with a binary variable, OLS regression is not suitable because the relationship between one or more predictors and the probability of a target outcome is nonlinear but assumes an S-shaped curve. This occurs because in a binary dependent variable the probabilities are bounded at values 0 and 1.

- In a linear regression – OLS predicted values are the mean of the target variable at the given values of the input variables;
- In a logistic regression predicted values are the probability of a particular level (or levels) of the target variable at the given values of the input variables.

CHAPTER 6. Analysis and Results

6.1. Binary logistic regression model

The analysis and model have been implemented primarily using STATA an econometric-statistical software capable of performing various functions: such as database management, statistical-econometric analysis, and graphical analysis. Other programs such as Microsoft Excel, gretl, and R were used as support in processing the data following collection via the survey, graphs, and tables.

Table 6. Binary logistic regression analyses compared.

	(1)	(2)
VARIABLES	**wtp**	**wtp**
fairness	0.171***	-
	-0.0467	-
shut	0.0164	-
	-0.0399	-
caught	0.00375	-
	-0.0322	-
penalty	-0.0294	-
	-0.0264	-
ethics	-	0.0854***
	-	-0.0194
risk aversion	-	-0.0189
	-	-0.0182
convenience	0.310***	0.310***
	-0.0454	-0.0452
mastery	0.0301	0.0321
	-0.0357	-0.0353
consumption	-0.0183	-0.0155
	-0.0468	-0.0461
discret. inc.	0.177**	0.187***
	-0.072	-0.0715
age	*yes*	*yes*
gender	*yes*	*yes*
education	*yes*	*yes*
Constant	-7.714***	-7.505***
	-1.368	-1.34
Observations	639	639

[5] [6]

[5] Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

[6] Table 6 shows the performance of two different binary logistic regressions: (1) performed using the four variables *fairness*, *shut*, *caught* and *penalty* left separate, (2) performed considering the variables *ethics* and *risk aversion* as the sum of two factors (*fairness* and *shut* for *ethics* and *caught* and *penalty* for *risk aversion*). The two different regressions are juxtaposed to observe and compare how the independent variables *xi* that remain unchanged (i.e., *convenience*, *mastery*, *consumption*, and *discretionary income*) change in significance across the two cases. It can be seen at first glance that the variable *x*

Table 7. Binary Logistic Regression.

Logistic regression		(1)
Log likelihood = -342.43024	**VARIABLES**	**wtp**
Number of obs = 639	ethics	0.0854***
		(0.0194)
LR chi2(28) = 177.93	risk	-0.0189
		(0.0182)
Prob > chi2 = 0.0000	convenience	0.310***
		(0.0452)
Pseudo R2 = 0.2062	mastery	0.0321
		(0.0353)
	consumption	-0.0155
		(0.0461)
	discret. inc.	0.187***
		(0.0715)
	age	*yes*
	gender	*yes*
	education	*yes*
	Constant	-7.505***
		(1.340)
	Observations	639
	Standard errors in	parentheses
	*** p<0.01, ** p<0.05,	* p<0.1.

In order to treat controls as factors, they have been run in STATA using the prefix "i.name of the factor", by doing so, STATA performs the dummy coding for the selected control variables. In all tables, controls' values have been replaced with the word "*yes*" meaning that controls have been applied to the model, but their values are not relevant to the model analyses of this study.

discretionary income is statically more significant to the model in the second (final) regression. This study analyzes and tests the regression comprehensive of the *ethics* and *risk aversion* variables shown in column (2).

6.2. Tests and Analyses

In evaluating a model there are two levels to pay attention to that are the overall fit of the model to the data (a), and the individual predictors in the model (b). After performing the Binary Logistic Regression, to test and interpret the results according to these two levels, analyses have been performed in order to determine whether the predictor is statistically significant associated with the dependent variable – i.e., the significance level (1), to determine the direction and effects of the predictor variables (2), and to determine how well the model fits – or does not fit – the data (Crowson, 2018).

6.2.1. Overall fit of the model to the data

Looking at Table 7, displaying the Binary Logistic Regression for this study, the values:

LR chi2(28) = 177.93
Prob > chi2 = 0.0000

represent the *likelihood ratio* **chi-squared test**; it tests whether the model that contains the full set of predictors represents a significant improvement in fit over the null model with no predictors. If this test is statistically significant then it would indicate that there is evidence of a good model fit in relation to the null model. As in this model, the *p-value* (equal to Prob > chi2 = 0.0000) is less than 0.01 or 0.05 thresholds, the null hypothesis is rejected, and the conclusion is that the current model exhibits significant improvement in fit over the baseline or null model (Crowson, 2018).

Pseudo R2 = 0.2062

Pseudo R^2 – or ρ^2 – is a measure for maximum likelihood estimation, based on the log-likelihoods (Seddighi, 2013). Some researchers consider ρ^2 as an analogy to the Least Squares R^2, paying attention though as it is not computed in the same way and does not yield the same results. Contrarily from the Least Squares Regression, Pseudo R^2 does not represent the proportion of variation in the dependent variable accounted for by the predictors; "*the values of ρ^2 tend to be considerably lower than those of the R^2 index and should not be judge by the standards for a "good fit" in ordinary regression analysis*" (McFadden, 1977). According to McFadden values of 0.2 to 0.4 for ρ^2 represent an excellent model fit, therefore, the value of ρ^2 in the binary logistic regression considered here is equal to 0.2062, representing an excellent result for the model fit.

Table 8. Binary logistic regression analysis, coefficients.

wtp	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
ethics	0.0854017	0.019428	4.4	0	0.0473236	0.1234799
risk	-0.018925	0.0181679	-1.04	0.298	-0.054533	0.0166837
convenience	0.3103888	0.0451767	6.87	0	0.2218441	0.3989335
mastery	0.0320953	0.0352716	0.91	0.363	-0.037036	0.1012262
consumption	-0.015463	0.0461274	-0.34	0.737	-0.105871	0.0749455
discretionary inc.	0.1865487	0.0715279	2.61	0.009	0.0463565	0.3267408
age	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
gender	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
educat.	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
_cons	-7.504536	1.340457	-5.6	0	-10.13178	-4.877289

6.2.2. The individual predictors in the model

Tables 7 and 8, show the predictor variables coefficients, standard errors, *z-values* (computed as a ratio of the regression coefficients to the standard errors), *p-values,* and 95% confidence interval. The results in a Binary Logistic Regression are interpreted similarly to the values obtained in a standard Least Squares Regression. In order to examine the association between the independent variables x_i and the dependent variable WTP, it is necessary to evaluate the *p-value*. The *p-value* represents the *"probability of obtaining results at least as extreme as the observed results of a statistical hypothesis test, assuming that the null hypothesis H_0 is correct"* (Investopedia, 2021). Moreover, the *p-value* determines the association and the significance of a predictor to the dependent variable; a *p-value* equal to or less than 0.05 is judged as statistically significant, and a *p-value* greater than 0.05 indicates that the predictor is not significant in the model. The *p-value* is therefore defined as a *probability* and its value is between 0 and 1. The smaller the *p-value*, the stronger the evidence in favor of the alternative hypothesis; this means that small p-values are evidence against the null hypothesis H_0. Vice versa, large p-values provide little evidence against H_0.

The null hypothesis states:

H_0: the coefficient of the term is equal to 0,

indicating that there is no association between the term and the response.

Looking at the Binary Logistic Regression analyzed in this study and shown in Table 8, the *p-value* of the term to the significance level is compared to evaluate the null hypothesis H_0; the null hypothesis can be *rejected* or *fail to be rejected*. The significance level is defined as the probability of rejecting or failing to reject the null hypothesis H_0. Significance levels are represented by the Greek letter α (alpha) and are 10% (when α=0.1), 5% (when α=0.05), and 1% (when α=0.01). If the hypothesis test yields a $p < \alpha$ – i.e., below the significance level – then the null hypothesis is rejected, and the result is statistically significant.

As displayed in table 7 and 8, looking at the *p-value* of the independent variables considered in this study, predictors *ethics*, *convenience*, and *discretionary income* are statistically significant at the level 0.01 (indicated by the three asterisks next to the coefficient value). The asterisks give the significance level representing the three-level of confidence 90%, 95% (*t-critic* value is 1.96), and 99% (*t-critic* value is 2.58) – respectively one, two, and three asterisks in the table. Therefore, the Binary Logistic Regression analysis shows that *ethics*, *convenience,* and *discretionary income* are significant for the model as they are statistically different from zero at the highest level of confidence (99%), while the other predictors do not show the significance for the model because their *p-values* are not smaller than the three alpha significance levels.

Table 9. Binary logistic regression analysis, odds ratio.

wtp	Odds Ratio	Std. Err.	z	P>z	[95% Conf.	Interval]
ethics	1.089155	0.0211601	4.4	0	1.048461	1.131427
risk	0.9812532	0.0178273	-1.04	0.298	0.9469271	1.016824
convenience	1.363955	0.061619	6.87	0	1.248377	1.490235
mastery	1.032616	0.036422	0.91	0.363	0.9636417	1.106527
consumption	0.9846564	0.0454196	-0.34	0.737	0.8995411	1.077825
discret. inc.	1.205083	0.0861971	2.61	0.009	1.047448	1.386442
age	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
gender	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
education	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
_cons	0.0005506	0.000738	-5.6	0	0.0000398	0.0076176

Confidence intervals

wtp	Coef.	[95% Conf.	Interval]
ethics	**0.0854017**	**0.0473236**	**0.1234799**
risk	-0.018925	-0.054533	0.0166837
convenience	**0.3103888**	**0.2218441**	**0.3989335**
mastery	0.0320953	-0.037036	0.1012262
consumption	-0.015463	-0.105871	0.0749455
discretionary income	**0.1865487**	**0.0463565**	**0.3267408**

The last two columns of Table 8 test the null hypothesis in terms of coefficients, as mentioned above:

H_0: The regression coefficient in the population is equal to 0.

By checking the values in the confidence interval, it can be determined whether the null value 0 falls within the interval or falls outside: as expected for the predictors *ethics*, *convenience,* and *discretionary income* the value 0 falls outside the interval, thus the null hypothesis is rejected.
Focusing on Table 9, displaying the odds ratios, the same results are expected as the same analysis is performed. In this situation the null hypothesis changes:

H_0: The odds ratio is equal to 1.

By checking again the values in the confidence interval:

wtp	Odds Ratio	[95% Conf.	Interval]
ethics	**1.089155**	**1.048461**	**1.131427**
risk	0.9812532	0.9469271	1.016824
convenience	**1.363955**	**1.248377**	**1.490235**
mastery	1.032616	0.9636417	1.106527
consumption	0.9846564	0.8995411	1.077825
discret. income	**1.205083**	**1.047448**	**1.386442**

for the predictors *ethics, convenience,* and *discretionary income* the null value 1 falls outside the confidence interval.

Table 8 displays the regression coefficients (in particular their slopes) and the tests of these coefficients. In a Logistic Regression Analysis, the coefficients represent the predicted change in

the *log-odds*; the predicted probability of group membership (target group). What is important to focus on is the predicted likelihood or probability of falling into the target group; this is determined by a ratio of two probabilities, probability of A which is "membership in the target group" over the probability of B which is "membership in the non-target group". This refers to the *odds* – a ratio of probabilities – the probability that one event will occur over the probability that another event will occur (assuming that these two events are mutually exclusive) (Crowson, 2018). In the Binary Logistic Regression Analysis considered in this dissertation, it is analyzed the probability of falling into the target group A *"willingness to subscribe to a streaming platform"* or the probability of falling into the non-target group B *"unwillingness to subscribe to a streaming platform"*. Moreover, the odds are determined by obtaining their *natural log*, reflecting the change in the *log-odds* for every one-unit increase on the predictor variable.

In a Logistic Regression Model it is necessary to look at the *log-odds* as, contrarily from the Least Squares Regression, the relationship between the predictors and the outcome variable is non-linear. The key concept is that:

- if the probability of the target event *equals* the probability of the non-target, the odds *equal* 1,
- if the probability of the target event is *greater* than the probability of the non-target event, the odds are *greater* than 1, and
- if the probability of the target event is *less* than the probability of the non-target event, the odds are *less* than 1.

The coefficients' slope of each independent variable represents indeed the predicted change in the probability of falling into the target group (relative to the reference group on the dependent variable) for a one-unit increase on the predictor (Seddighi, 2013). The coefficients' slope gives the partial effects of each x_i on the response probability by determining its direction. In Chapter 5 are reported the hypotheses about the positive or negative relationship between the predictors x_i and the dependent variable WTP. A positive regression coefficient represents a positive relationship among the x predictor and y – i.e., the probability of falling into the target group increases as a result of increases on the predictor variable, vice versa a negative sign shows a negative or inverse relation among x and y – i.e., the probability of target membership decreases with increases on the predictor. Our results provide empirical support for Hypotheses 1, 2, and 6 and indicate that *ethics*, as well as *discretionary income* and *convenience*, increases the likelihood of paid subscriptions for digital music. The coefficient for these variables is positive and significant at $p<0.01$. Conversely, no empirical evidence is found in support of Hypotheses 3, 4, and 5 indicating that there is no

statistically significant relationship between independent variables *risk aversion*, *mastery of hardware use* and *consumption through other channels,* and WTP for digital music. Looking at the variables that are statistically significant to the model, our results show that the signs of their coefficients correspond with assumptions previously estimated; precisely variables *ethics*, *convenience,* and *discretionary income* present a positive sign +. showing that:
Specifically, a positive value in *ethics* (*convenience* and *discretionary income*) indicates a positive relation to the likelihood of falling into the target group – so the likelihood of paying for a streaming platform subscription *WTP*, indeed higher levels of *ethics* yield to a greater likelihood to *subscribe* while lower levels of *ethics* yield to less probability to *subscribe*.

As the odds represent a ratio of probabilities, probability of falling into the target group A over the probability of falling into the non-target group B, the odds ratios calculated in Table 9 are reflecting the changes in odds for every one-unit increase on the predictor variable.
Table 9 shows the value of the *odds ratio* for each predictor x_i. The *odds ratio* is used to determine the effect of a predictor on the response variable; specifically, if the odds ratio is greater than 1 it indicates that the event is *more* likely to occur as the predictor increases while an odds ratio that is less than 1 indicates that the event is *less* likely to occur as the predictor increases. Considering variable *ethics*: for every one-unit increase on *ethics* the odds of WTP for digital music change by a factor of 1.089, i.e., meaning that the odds are increasing (in parallel, a change by a factor of 1.364 for *convenience* and 1.205 for *discretionary income*). In contrast, for a value of an odds ratio smaller than 1*;* for every one-unit increase on the predictor, the odds of the response change by a factor of the value of the odds ratio, meaning a decrease in odds.

6.2.3. Post estimation analyses

In order to evaluate the goodness of fit of the model (together with the chi-squared test previously performed), two additional tests are performed: the *Hosmer-Lemeshow* test (1), and the *Pearson chi-square* test (2).

Logistic model for wtp, goodness-of-fit test (1)

number of observations = 639
number of groups = 10
Hosmer-Lemeshow chi2(8) = 5.38
Prob > chi2 = 0.7165

Logistic model for wtp, goodness-of-fit test (2)

number of observations = 639
number of covariate patterns = 639
Pearson chi2(610) = 638.82
Prob > chi2 = 0.2028

Unlike for the chi-squared test above where a significant *p-value* is an indicator of good model fit, in the Hosmer-Lemeshow test the focus is on *non-significance* (Crowson, 2018): by checking the values on the Hosmer-Lemeshow test (1), the *p-value* equals 0.7165 – therefore it is not significant meaning that it is an indicator of good model fit. Likewise, a non-significant test result in the Pearson chi-square test (2), *p-value* equal to 0.2028, is an indicator of good model fit. Moreover, these results can be interpreted in reverse; as the values of the goodness-of-fit tests are all greater than the significance level of 0.05, it indicates that there is not enough evidence to conclude that the model does not fit the data.

Another analysis to test the good fit of the model is the *Classification* Table or *Sensitivity* Test (Table 10) which displays the classification results that are determined by the model which generates the predicted probabilities for group membership or no-membership. Based on the predicted probabilities, an estimation can be generated as to whether an individual falls into group 0 *WTP* or group 1 *"unwillingness to pay"*. The *Classification* Table compares the observed group membership and the group membership that is predicted based on the model.

Table 10. Classification Table.

	--- True ---		
Classified	D	~D	Total
+	**320**	108	428
-	60	**151**	211
Total	**380**	**259**	639

Classified + if predicted Pr(D)>= 0.5

True D defined as wtp != 0

Sensitivity		**Pr(+ D)**	**84.21%**
Specificity		**Pr(- ~D)**	**58.30%**
Positive predictive value	Pr(D +)	74.77%	
Negative predictive value	Pr(~D -)	71.56%	
False + rate for true ~D		Pr(+~D)	41.70%
False - rate for true D		Pr(- D)	15.79%
False + rate for classified +	Pr(~D +)	25.23%	
Correctly classified			**73.71%**

From the *Classification* Table above, 380 cases are observed to fall into the WTP group whereas 259 cases into the "*unwillingness to pay"* group. 320 out of the total of 380 cases of the WTP group are predicted correctly by the model to fall into that category – the accuracy rate for that group is 84.21% (relatively good/high percentage). While 151 over 259 individuals are predicted by the model to express intention *not to subscribe* – represented by the percentage 58.30% (approximately half of the total). Summarizing 471 cases (sum of 320 and 151) are correctly classified based on the model, a total accuracy rate of about 73.71%, which indicates a relatively good model estimation.

6.3. Results

6.3.1. Regression analysis results

In this dissertation, a Binary Logistic Regression (or Logit Model) is used to analyze the data collected. Initially, two different regressions performed on STATA with the use of different variables have been compared; the first regression has been carried out keeping the variables *fairness, shut, caught,* and *penalty* separated (and therefore not considering them as a sum of variables) (1), and a second regression (final) which included a total of 6 independent variables *x,* of which 2, *ethics*

and *risk aversion,* are the sum of two factors each (2)[7]. The two different regressions completed showed, as predicted by the previously made assumptions, that the second regression demonstrates more consistent results with the model. Moreover, the significance levels of the variables x_i associated with the response variable *y*, are higher in the second – final regression.

Several analyses have been performed at both the aggregate and specific levels. For each independent variable, the significance level for the model has been analyzed, in order to assess the association with the dependent variable *y* WTP for music streaming platforms. From the analyses that were performed, it can be concluded that 3 out of 6 variables – *ethics*, *convenience,* and *discretionary income* – are significant for the model and that their direction – slope matches with the expectations established in the initial hypotheses. In evaluating how well the model fits the data, the analyses that are completed have shown a good fit of the model to the data. Additional verification was carried out to establish whether the model was able to predict independently the correct cases of membership to the target group or to the non-target group (*sensitivity test*) and that resulted in a good rate of estimation by the model. For the variables *risk aversion, mastery,* and *consumption* no further testing has been performed as these variables were found to be non-significant for the model.

In relation to previous studies in the literature, this analysis also confirms that ethical factors related to the belief that it is wrong to engage in pro-piracy activities result in an inclination to pay for streaming platforms. Similarly, as previously analyzed in the literature, this research also confirms that income – particularly discretionary income – is influential in users' WTP for subscriptions to digital platforms.

Table 11. Summary of hypotheses testing results.

HYPOTHESIS	(1) **conclusion**	(2) **conclusion**
H1	*yes*	*supported*
H2	*yes*	*supported*
H3	no	not supported
H4	no	not supported
H5	no	not supported
H6	*yes*	*supported*

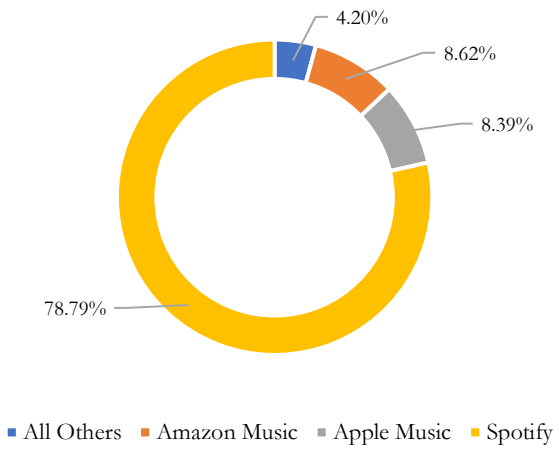
[7] Conception Model, Chapter 5

6.3.2. General study results

In the survey distributed, in addition to the questions about the variables used to build the model and the binary logistic regression, questions were included at the descriptive and information level, with the intent to collect data on the habits and preferences of consumers regarding music streaming platforms.

Specifically, respondents have been asked whether they were currently paying for a streaming platform and if this was the case, for which one: it resulted that 640 out of 859 respondents (corresponding to 74.51%) pay for a music streaming platform subscription. Specifically, 78.79% pay for a subscription to Spotify, followed by Amazon Music and Apple Music with a percentage of 8.62% and 8.39%. All other platforms representing 4.20%.

Figure 5. Streaming platforms subscriptions.

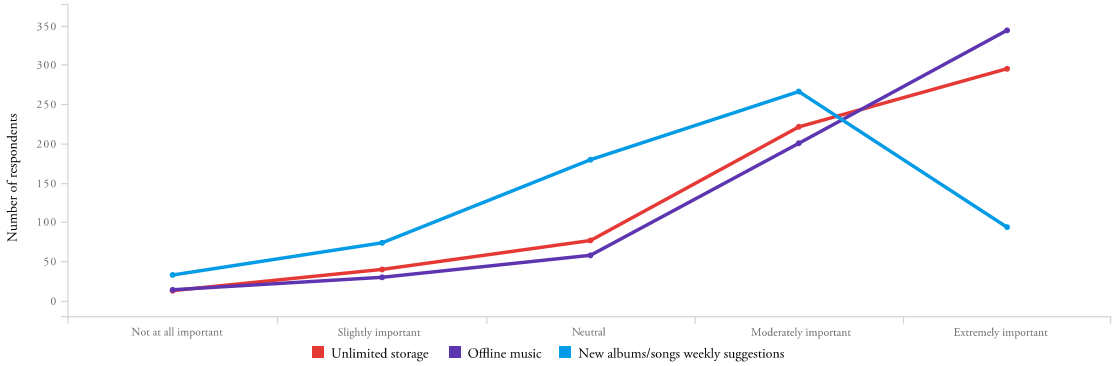


Moreover, respondents were asked how much value they place on certain specific characteristics and features of music streaming platforms, Figure 6 shows that:

- 80.25% of respondents consider it important to have access to a music streaming platform that offers *unlimited storage* for music.
- 84.6% of respondents consider it important to have access to *music offline.*
- 55.83% of respondents, to a lesser extent than above, rate important to receive *new albums/songs weekly suggestions from the platform.*[8]

[8] Importance percentages are expressed as a sum of responses "Moderately important" and "Extremely important" on a 5-point Likert Scale.

Figure 6. Music streaming platforms' features.



In conclusion, in the case a person is paying for a subscription to a music streaming platform, based on personal experience, questions were asked about: to which extent a consumer is satisfied or not with the subscription/use of the music streaming platform (1) and how likely is a person to recommend the substitution of a P2P file-sharing website with a subscription to a music streaming platform (2).

The results collected show that 91.89% of respondents are satisfied[9] with the subscription they are paying and 68.59% are likely[10] to recommend the substitution of illegal platforms with a legal subscription.

[9] Satisfaction percentages are expressed as a sum of responses "Mostly satisfied" and "Completely satisfied" on a 5-point Likert Scale.

[10] Like hood percentages are expressed as a sum of responses "Likely" and "Extremely Likely" on a 5-point Likert Scale.

Limitations

One potential limitation of this study is based on the data collected; our survey data is based solely on a sample of people between the ages of 18 and 35 years old. This choice is due to the fact that this study wanted to focus on a precise age segment, i.e. it excludes professional seniors (over 35), younger people (underage), and older people (retirees).

83% of respondents are Italian citizens, for this reason, the sample is not large enough to include other related markets, particularly those people of other nationalities. It would therefore be interesting to expand this analysis to study the same phenomenon outside Italy.

For privacy reasons, in this study respondents were not asked directly if they had used pirate sites or illegal downloads. The data collection is based on their perception of these P2P platforms and not directly on their use. For this reason, a limitation of this analysis is that it does not directly analyze the percentage of people who use piracy as a method of acquiring digital music.

This analysis investigates the willingness to pay for a subscription to a digital platform; it does not consider that segment of consumers who legally use streaming platforms in free mode. This limitation can be overcome by analyzing at an aggregate level the users of legal music streaming platforms in free mode and in paid subscription mode.

Conclusion

This dissertation aims to investigate the variables that influence individuals' willingness to pay for digital music by studying in more detail their willingness for subscriptions to music streaming platforms. Through a binary logistic regression analysis, six hypotheses, representing six variables, were analyzed to test their possible influence on users' WTP for digital music. Following data collection through a survey and a logistic regression analysis performed on the software STATA, it was found that out of the six hypotheses formulated at the beginning of this study, three were confirmed to be significant to the analysis conducted.

The variable *ethics* was studied as a sum of the extent to which a person considers it fair to pay for digital music (*fairness*) and the extent to which a person believes that illegal download platforms should be shut down (*shut*); *ethics* was found to be influential to users' WTP for a subscription to music streaming platforms. Similarly, building on studies conducted in the prior literature on the topic, this study confirmed that the discretionary income of surveyed individuals has a positive effect on their WTP for streaming music subscriptions. In fact, *discretionary income* as an analyzed variable was found to be statistically significant in our analysis. Moreover, it was hypothesized that the *convenience* factor, intended as the convenience in the use of streaming platforms, thanks to their tools and features, can be influential in the choice to pay for these platforms themselves. Our analysis also supported this hypothesis as a result of a significant *convenience* variable.

On the contrary, the hypotheses related to a possible link between *risk aversion*, the *consumption* of music through other channels – i.e., channels other than streaming – and *mastery* in the use of technologies were not significant and relevant for the users' WTP for digital music. With the hope that this research will further advance the understanding of the factors influencing people's willingness to pay for art-related digital music products, future studies are encouraged to investigate this topic further in order to increasingly reduce the use of piracy.

Bibliography

Aguiar, L., Waldfogel, J., (2015) "Streaming reaches flood stage: Does Spotify stimulate or de-press music sales?". National Bureau of Economic Research.

Aguiar, Luis and Martens, Bertin. (2016) "Digital music consumption on the Internet: Evidence from clickstream data". Information Economics and Policy, Vol. 34.

Aguiar, Luis. (2017) "Let the music play? Free streaming and its effects on digital music consumption". Information Economics and Policy.

Ahmadi, M., (2015) "Key Factors Affecting Consumer Music Procurement Behavior (Observing Music Sites)".

Al-Rafee, S., Cronan, T.P., (2006) "Digital piracy: factors that influence attitude toward behavior".

Alexander, Peter J. (2002) "Peer-to-Peer File Sharing: The Case of the Music Recording Industry". Review of Industrial Organization.

Alleyne, Philmore & Soleyn, Sherlexis & Harris, Terry. (2015) "Predicting Accounting Students' Intentions to Engage in Software and Music Piracy". Journal of Academic Ethics.

Álvarez Vázquez, R. (2017) "The music industry in the dawn of the 21st century Networking for a thriving music industry".

Andersen, B., Frenz, M., (2010) "Don't blame the P2P file-sharers: the impact of free music downloads on the purchase of music CDs in Canada", J. Evol. Econ. 20 (5), 715–740.

Anderson, S., & Coate, S. (2005). "Market Provision of Broadcasting: A Welfare Analysis". The Review of Economic Studies, 72(4), 947-972.

Andres, A., (2006) "Software piracy and income inequality". Applied Economics Letters. 13. 101-105.

Armstrong, Mark & Weeds, Helen. (2007) "Public Service Broadcasting in the Digital World".

Baird, Katharine, (2018) "Advertising Appeals and Willingness to Pay for a Music Streaming Service" Undergraduate Honors Thesis Collection. 442.

Banerjee, D. S. (2003) "Software Piracy: A Strategic Analysis and Policy Instruments." International Journal of Industrial Organization.

Bartsch, K. (2017) "The Napster Moment: Access and Innovation in Academic Publishing", 343 – 348.

Becker, G. (1968) "Crime and Punishment: An Economic Approach." Journal of Political Economy.

Bekir, I., (2017), "Corruption, Income and Piracy. An empirical analysis". Department of Economy, University of Sousse.

Belleflamme, Paul and Peitz, Martin, (2010) "Digital Piracy: Theory". CESifo Working Paper Series No. 3222.

Ben-Shahar, D., and A. Jacob. (2004) "Selective Enforcement of Copyright as an Optimal Monopolistic Behavior".

Benayoune F., Lancieri L. (2004) "Models of Cooperation in Peer-to-Peer Networks – A Survey". In: Freire M.M., Chemouil P., Lorenz P., Gravey A. (eds) Universal Multiservice Networks. ECUMN 2004. Lecture Notes in Computer Science, vol 3262. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-30197-4_33.

Bhattacharjee, S., Gopal, R., Sanders, G., (2003) "Digital music and online sharing: software piracy 2.0?".

Bhattacharjee, S., R. Gopal, K. Lertwachara, and J. Marsden, (2006) "Impact of Legal Threats on Online Music Sharing Activity: An Analysis of Music Indus- try Legal Actions." Journal of Law and Economics.

Boccanegra, C. (2015) "Il Crowdfunding e l'industria musicale".

Boldrin, M., and D. K. Levine. (2002) "The Case Against Intellectual Property." American Economic Review.

Borja, K., Dieringer, S., Daw, J., (2015) "The effect of music streaming services on music piracy among college students".

Bowie, N., (2005) "Digital rights and wrongs: intellectual property in the information age".

Chen, C., Chen, C., and Yeh, C., (2010) "Determinants of Software Piracy: Evidence from Far East Countries", Journal of International and Global Economic Studies.

Cheng, H. K., R. Sims, and H. Teegen. (1997) "To Purchase or to Pirate Software: An Empirical Study." Journal of Management Information Systems.

Chiang, E. P. and Assane, D. (2009) "Estimating the willingness to pay for digital music", Contemporary Economic Policy, 27(4), pp. 512–522.

Chiang, E. P., and D. Assane. (2007) "Determinants of Music Copyright Violations on the University Campus." Journal of Cultural Economics.

Chiariglione, L. (1998) "Impact of MPEG standards on multimedia industry". Proceedings of the IEEE, 86(6), 1222–1227. https://doi.org/10.1109/5.687836.

Christensen, C., & Euchner, J. (2020) Managing Disruption. Research-Technology Management, 63(3), 49–54.

Cohen, E., and L. Cornwell. (1989) "College Students Believe Piracy is Acceptable".

Conner, K. R., and R. P. Rumelt. (1991) "Software Piracy: An Analysis of Protection Strategies." Management Science.

Crowson, H. & Brandes, Joyce. (2017). "Differentiating Between Donald Trump and Hillary Clinton Voters Using Facets of Right-Wing Authoritarianism and Social-Dominance Orientation: A Brief Report." Psychological Reports.

Cunningham, Stuart D. (2002) "From cultural to creative industries: Theory, industry and policy implications". Media International Australia Incorporating Culture and Policy: Quarterly Journal of Media Research and Resources.

Danaher, B., Smith, M. D., Teland, R., "Piracy and copyright enforcement mechanisms". National Bureau of Economic Research.

Dang Nguyen, G., Dejean, S. and Moreau, F. (2012) "Are Streaming and Other Music Consumption Modes Substitutes or Complements?", SSRN Electronic Journal, pp. 1–22.

Delikan, M. D. (2010) "Changing Consumption Behavior of Net Generation and the Adoption of Streaming Music Services", (June), pp. 1–53.

Downs, K. and Srivastava, S. (2020) "How streaming is changing the music industry | Deloitte Insights", Deloitte Insights, pp. 1–5.

Eiriz, Vasco and Leite, Filipe P. (2017) "The digital distribution of music and its impact on the business models of independent musicians". The Service Industries Journal.

Field, A. (2018) "Discovering statistics using IBM SPSS statistics" (5th ed).

Flew, Terry and Cunningham, Stuart. (2010) "Creative Industries after the First Decade of Debate". The Information Society.

Fly, B. (2016) "How Does Music Consumption Impact the Music Industry and Benefit Artists?".

Gayer, A., and O. Shy. (2003) "Copyright Protection and Hardware Taxation." Information Economics and Policy.

Gopal, R. D., and G. L. Sanders. (1998) "International Software Piracy: Analysis of Key Issues and Impacts." Information Systems Research.

Gramolini, F. (2015) "L' innovazione digitale nell' industria musicale: il player Apple". Grassini, S. (2012) "La creazione e l'appropriazione di valore nell' industria discografica".

Gupta, P. B., Gould, S. J., & Pola, B. (2004) "To pirate or not to pirate": A comparative study of the ethical versus other influences on the consumer's software acquisition-mode decision. Journal of Business Ethics, 55(3), 255-274.

Hardré, P. L., Crowson, H. M., & Xie, K. (2012). "Examining Contexts-of-Use for Web-Based and Paper-Based Questionnaires." Educational and Psychological Measurement, 72(6), 1015–1038.

Hracs, Brian J. (2012) "A Creative Industry in Transition: The Rise of Digitally Driven Independent Music Production". Growth and Change: a Journal of urban and regional policy.

Hracs, Brian J. (2015) "Cultural Intermediaries in the Digital Age: The Case of Independent Musicians and Managers in Toronto". Regional Studies.

Johnson, N., and S. Kotz. (1994) Continuous Univariate Distribution. 2nd ed., Vol. 2. New York: John Wiley & Sons.

Keller, M. (2019, March 4) "Sweden Pulls the Plug on Pirate Bay". Los Angeles Times. https://www.latimes.com/archives/la-xpm-2006-jun-01-fi-piratebay1-story.html.

King, S. P., and R. Lampe. (2003) "Network Externalities, Price Discrimination and Profitable Piracy."

Ku, R. S. R. (2002) "The Creative Destruction of Copyright: Napster and the New Economics of Digital Technology", The University of Chicago Law Review, 69(1), pp. 263–324.

Kwong, K.K., Yau, O.H.M., Lee, J.S.Y. et al. (2003) "The Effects of Attitudinal and Demo-graphic Factors on Intention to Buy Pirated CDs: The Case of Chinese Consumers". Journal of Business Ethics 47, 223–235. https://doi.org/10.1023/A:1026269003472

Kyper, E., Prante, G., and Schimmoeller, L., (2016) "Evaluating the role of income in national software piracy". Advances in Business Research.

Larabi, Chaker & Rosselli, Vincent & Fernandez-Maloigne, Christine. (2009) "Image Compression Standards and their HVS-Based Optimization: Application to Digital Cinema Profile of JPEG 2000".

Leocata, M. A., (2017) "Il mercato musicale nell'era dei social network".

Lertwachara, K., & Ayanso, A., & Molla, A., (2007) "An Analysis of the Effectiveness of Law Enforcement Against Online Music Piracy".

Liou, Shyhnan and Yang, Chia Han. (2016) "Innovation in Cultural and Creative Industries". World Academy of Science, Engineering and Technology.

Lomax, R.G., & Hahs-Vaughn (2012) "An introduction to statistical concepts" (3rd ed).

Longo, A. (2014) "Il mercato della musica digitale. Il case study su Spotify", pp. 1–79.

Lysonski, Steven & Durvasula, Srinivas. (2008) "Digital Piracy of MP3s: Consumer and Ethical Predispositions". Marketing Faculty Research and Publications. 25. 10.1108/07363760810870662.

Marron, D. B., and D. G. Steel. (2000) "Which Countries Protect Intellectual Property? The Case of Software Piracy." Economic Inquiry.

Moore, F. (2020) "IFPI Global Music Report 2019", (April 2019), pp. 2019–2020.

Mulligan, M., "Recorded Music Revenues Hit $21.5 Billion in 2019", https://www.midiaresearch.com/blog/recorded-music-revenues-hit-215-billion-in-2019.

Naveed, K., Watanabe, C. and Neittaanmäki, P. (2017) "Co-evolution between streaming and live music leads a way to the sustainable growth of music industry – Lessons from the US experiences", Technology in Society, 50, pp. 1–19.

Nyaw, M., and I. Ng. (1994) "A Comparative Analysis of Ethical Beliefs: A Four Country Study." Journal of Business Ethics.

Oberholzer-Gee, F., and K. Strumpf. (2007) "The Effect of File Sharing on Record Sales: An Empirical Analysis." Journal of Political Economy.

Osborne, J.W. (2015) "Best practices in logistic regression".

Peitz M., and Valletti T., (2008) "Content and advertising in the media: Pay-tv versus free-to-air", International Journal of Industrial Organization, vol. 26, issue 4, 949-965.

Peitz, M. (2004) "A Strategic Approach to Software Protection: Comment." Journal of Economics & Management Strategy.

Peron, E., & Peron, E. (2016, May 16). "L'evoluzione dell'industria musicale nell'era digitale", This MARKETERs Life. https://www.thismarketerslife.it/digital/dal-compact-disc-allo-streaming-levoluzione-dellindustria-musicale-nellera-digitale/.

Pham, Q. T., Dang, N. M., & Nguyen, D. T., (2020) "Factors Affecting on the Digital Piracy Behavior: An Empirical Study in Vietnam". Journal of theoretical and applied electronic commerce research, 15(2), 122-135.

Pingitore, P. (2016) "Digital Music nell'industria discografica: il ruolo delle Major labels e di Spotify".

Pittman, E. W. (2016) "Are Music Streaming Services Healthy for the Recorded Music Indus-try?", p. 29.

Pituch, K.A., & Stevens, J.A. (2016) "Applied multivariate statistics for the social sciences" (6th ed).

Prey, R. (2018) "Nothing personal: algorithmic individuation on music streaming platforms".

Seddighi (2013) "Introductory Econometrics", Introductory Econometrics".

SH Bae, JP Choi, (2006) "A model of piracy", Information Economics and Policy.

Shy, O., and J. F. Thisse. (1999) "A Strategic Approach to Software Protection." Journal of Economics & Management Strategy.

Silva, F., and G. B. Ramello. (2000) "Sound Recording Market: The Ambiguous Case of Copy-right and Piracy." Industrial and Corporate Change.

Sims, R., H. K. Cheng, and H. Teegen. (1996) "Toward a Profile of Student Software Piraters." Journal of Business Ethics.

Sinclair, Gary & Green, Todd. (2015) "Download or stream? Steal or buy? Developing a typology of today's music consumer". Journal of Consumer Behaviour. 15. 10.1002/cb.1526.

Slive, J., and D. Bernhardt. (1998) "Pirated for Profit." Canadian Journal of Economics.

Smith, T.J., & McKenna, C.M. (2013) "A comparison of logistic regression pseudo R2. Multiple Linear Regression Viewpoints", 39, 17-26.

Stolpe, M. (2000) "Protection against Software Piracy: A Study of Technology Adoption for the Enforcement of Intellectual Property Rights." Economics of Innovation and New Technology. study".

Sudler, H. (2013) "Effectiveness of anti-piracy technology: Finding appropriate solutions for evolving online piracy", Business Horizons, 56(2), pp. 149–157.

Suduc, A.-M., Bizoi, M., & Filip, F. G. (2009). Ethical Aspects on Software Piracy and Information and Communication Technologies Misuse. IFAC Proceedings Volumes, 42(25), 30–35.

Swinyard, W., H. Rinne, and A. Kau. (1990) "The Morality of Software Piracy: A Cross-Cultural Analysis." Jour-nal of Business Ethics.

Tabachnick, B.G., & Fidell, L.S. (2013). "Using multivariate statistics" (6th ed.).

Thomes, T. P. (2013) "An economic analysis of online streaming music services", Information Economics and Policy, 25(2), pp. 81–91.

Throsby, D. (2002), "The music industry in the new millennium: Global and local perspectives".

Varian, H. (2005) "Copying and Copyright." Journal of Economic Perspectives.

Veitch, R. W. D., Constantiou, I., (2011) "To Pirate or purchase? A model of individual decisions for digital product acquisition", Copenhagen Business School.

Volgsten, U., 2019. A technology and its vicissitudes: playing the gramophone in Sweden 1903–1945. Popular Music, 38(2), pp.219-236.

Waldfogel, Joel. (2012) "Music Piracy and Its Effect on Demand, Supply, and Welfare", Innovation Policy and the Economy.

Wang, F., Zhang, H., & Ouyang, M. (2005) "Software Piracy and Ethical Decision-Making Behavior of Chinese Consumers". Journal of Comparative International Management, 8(2).

Yoon, K. (2002) "The Optimal Level of Copyright Protection." Information Economics and Policy.

Zentner, A. (2006) "Measuring the Effect of File Sharing on Music Purchases." Journal of Law and Economics.

Sitography

Amazon.it. 2021. Amazon.it: Prime Music. [online] Available at: <https://www.amazon.it/music/prime> [Accessed November 2020].

Apple (Italia). 2021. Apple Music. [online] Available at: <https://www.apple.com/it/apple-music/> [Accessed November 2020].

Billboard. 2021. Billboard - Music Charts, News, Photos & Video. [online] Available at: <https://www.billboard.com/> [Accessed July 2020].

Cercle.io. 2021. Cercle | Timeless Music. [online] Available at: <https://www.cercle.io> [Accessed November 2020].

Deezer. 2021. Deezer | Ascolta musica | Piattaforma di streaming musicale online. [online] Available at: <https://www.deezer.com/it/> [Accessed November 2020].

Dittomusic.com. 2021. [online] Available at: <https://dittomusic.com/en/?gclid=EAIaIQobChMIo9mHv_TT7wIVAgGLCh20ugqyEAAYASAAEgLbBvD_BwE> [Accessed September 2020].

Donà, N., 2017. Dai Dischi al Digitale: La Strana Storia Dei Formati Musicali | LANDR Blog. [online] LANDR Blog. Available at: <https://blog.landr.com/it/storia-dei-formati-musicali/> [Accessed December 2020].

IFPI. 2021. IFPI — Representing the recording industry worldwide. [online] Available at: <https://www.ifpi.org/> [Accessed 28 March 2020].

Investopedia. 2021. Investopedia. [online] Available at: <https://www.investopedia.com/> [Accessed September 2020].

Napster. 2021. Home | Napster. [online] Available at: <https://it.napster.com/> [Accessed December 2020].

Pandora.com. 2021. Pandora. [online] Available at: <https://www.pandora.com/> [Accessed November 2020].

Resident Advisor. 2021. RA / Electronic music online. [online] Available at: <https://ra.co> [Accessed January 2021].

RIAA. 2021. Home - RIAA. [online] Available at: <https://www.riaa.com/> [Accessed November 2020].

SoundCloud. 2021. SoundCloud – Listen to free music and podcasts on SoundCloud. [online] Available at: <https://soundcloud.com/> [Accessed November 2020].

Spotify.com. 2021. Listening is Everything. [online] Available at: <https://www.spotify.com/it/> [Accessed November 2020].

Support.minitab.com. 2021. Interpret the key results for Fit Binary Logistic Model - Minitab. [online] Available at: <https://support.minitab.com/en-us/minitab/19/help-and-how-to/statistical-modeling/regression/how-to/fit-binary-logistic-model/interpret-the-results/key-results/> [Accessed November 2020].

TIDAL - High Fidelity Music Streaming. 2021. TIDAL - High Fidelity Music Streaming. [online] Available at: <https://tidal.com/> [Accessed November 2020].

Youtube.com. 2021. [online] Available at: <https://www.youtube.com/musicpremium> [Accessed October 2020].

Appendix

Table 12. Binary logistic regression analysis, coefficients (1).

wtp	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
fairness	0.17	.046747	3.66	0.000	.0794271	0.26
shut	0.02	.0399495	0.41	0.681	-.0618663	0.09
caught	0.00	.0321801	0.12	0.907	-.0593208	0.07
penalty	-0.03	.0264249	-1.11	0.266	-.0811815	0.02
convenience	0.31	.0453612	6.83	0.000	.220746	0.40
mastery	0.03	.0356806	0.84	0.399	-.039838	0.10
consumption	-0.02	.0468413	-0.39	0.696	-.1101084	0.07
discretionary inc.	0.18	.0719971	2.45	0.014	.0356337	0.32
age	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
gender	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
educat.	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
_cons	-7.714052	1.367581	-5.64	0.000	-10.39446	-5.033644

Table 13. Binary logistic regression analysis, odds ratio (1).

wtp	Odds Ratio	Std. Err.	z	P>z	[95% Conf.	Interval]
fairness	1.18655	0.0554677	3.66	0	1.082667	1.3004
shut	1.016569	0.0406115	0.41	0.681	0.9400085	1.099365
caught	1.003758	0.032301	0.12	0.907	0.9424044	1.069106
penalty	0.971038	0.0256596	-1.11	0.266	0.9220263	1.022655
convenience	1.362951	0.0618251	6.83	0	1.247007	1.489676
mastery	1.030552	0.0367707	0.84	0.399	0.9609451	1.105201
consumption	0.9818654	0.0459919	-0.39	0.696	0.8957371	1.076275
discret. inc.	1.193327	0.0859161	2.45	0.014	1.036276	1.37418
age	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
gender	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
educat.	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
_cons	0.0004465	0.0006106	-5.64	0	0.0000306	0.006515

Table 14. Qualtrics survey.

SCREENING QUESTIONS	**POSSIBLE ANSWERS**				
How often do you listen to music in a digital format?	Never	Rarely	Sometimes	Often	Always
Are you aware of the existence of one or more of the following illegal P2P file- sharing platforms?	eMule	BitTorrent	Napster	YouTube mp3 converter	Spotify Premium Crack
Are you aware that downloading music through P2P is illegal and prosecuted by law?	Yes	No			
SURVEY QUESTIONS					
How old are you?	(*choose*) 18-35				
With which gender identity do you most identify?	Male	Female	Other	Prefer not to disclose	
What is your most recent educational achievement that you have successfully completed?	Less than High School Diploma or Professional Certificate	High School Diploma	Bachelor's degree	Master's degree	PhD
What is your (monthly) discretionary income?	200 EUR or less	200 EUR - 500 EUR	500 EUR - 800 EUR	800 EUR - 1000 EUR	1000 EUR or more
What is your nationality?	(*choose*)				
Are you currently enrolled as a student?	Yes	No			
Are you currently working?	Yes full-time	Yes part-time	No		
Is your study major or work closely related to the field of technology (i.e. computer science, programming, IT support, engineering, economics, business, mathematics, design, architecture, or similar)?	Yes	No			
Does your current occupation require you to use technology (i.e. computer, tablet)?	Never	Rarely	Sometimes	Often	Always
Think about your recent experience of listening to music in a digital format. Please indicate how much you either agree or disagree with each statement below.					
Piracy is unfair towards musicians.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Because of piracy musicians are not adequately compensated for their effort.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Downloading music illegally is an act of stealing.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
People should pay for digital music.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
P2P illegal platforms should be shut down because they hinder the sustainability of the music industry.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
P2P illegal platforms should be shut down because they are not safe.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree

P2P illegal platforms should be shut down because they are unethical to the work of artists and record labels.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
P2P illegal platforms should be closed because they are against the law.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree

Please indicate how likely do you think the situations below would happen.

Getting caught for downloading music through P2P platforms.	Extremely unlikely	Unlikely	Neutral	Likely	Extremely likely
Getting caught for uploading music on P2P platforms.	Extremely unlikely	Unlikely	Neutral	Likely	Extremely likely
Getting caught by operating systems (iOS, Android) for downloading Premium cracked versions of music applications.	Extremely unlikely	Unlikely	Neutral	Likely	Extremely likely
Getting caught to be in possession of illegally downloaded music material.	Extremely unlikely	Unlikely	Neutral	Likely	Extremely likely

How concerned are you with the following statements?

I will pay a large fine for illegal downloading.	Not at all concerned	Slightly concerned	Somewhat concerned	Moderately concerned	Extremely concerned
I will go to prison for illegal downloading.	Not at all concerned	Slightly concerned	Somewhat concerned	Moderately concerned	Extremely concerned
Illegal downloading might affect my reputation negatively.	Not at all concerned	Slightly concerned	Somewhat concerned	Moderately concerned	Extremely concerned
Illegal downloading might affect my career negatively.	Not at all concerned	Slightly concerned	Somewhat concerned	Moderately concerned	Extremely concerned

Think of your recent experience of using streaming music platforms – such as Spotify, Deezer, Apple Music, Pandora, etc. – and please indicate how much you either agree or disagree with each statement below.

It is easy to use a music streaming platform.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I like having all my playlists and songs stored in a single place.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I like that streaming music platforms are synced with my other social media and/or apps.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I like automatic suggestions for new tracks and artists based on my preferences.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I enjoy using new technologies.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I feel confident using technology devices and hardware (i.e. smartphone, tablet, laptop, PC).	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I feel confident in using new apps and software.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I believe my ability to use technology is above average.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree

How often do you listen to the radio?	Never	Rarely	Sometimes	Often	Always

How often do you go to live concerts?	Never	Rarely	Sometimes	Often	Always
How often do you purchase physical music (non digital forms of music) e.g. CDs, Vinyl?	Never	Rarely	Sometimes	Often	Always
How often do you purchase tracks on iTunes?	Never	Rarely	Sometimes	Often	Always

How much value do you place on the following features?

Unlimited storage	Not at all important	Slightly important	Neutral	Moderately important	Extremely important
Offline music	Not at all important	Slightly important	Neutral	Moderately important	Extremely important
New albums/songs weekly suggestions	Not at all important	Slightly important	Neutral	Moderately important	Extremely important

Are you paying for a monthly subscription to a music streaming platform?	Yes	No			
If yes, which one?	(choose) Spotify - Deezer - Amazon Music - Apple Music - Tidal - YouTube Music - SoundCloud - Other				
How satisfied are you with your subscription/use of the music streaming platform?	Completely dissatisfied	Mostly dissatisfied	Neither satisfied nor dissatisfied	Mostly satisfied	Completel y satisfied
How likely are you to recommend the substitution of a P2P file-sharing website with a subscription to a music streaming platform based on your experience?	Extremely unlikely	Unlikely	Neutral	Likely	Extremely likely