UNIVERSITA' DELLA CALABRIA

Dipartimento di Economia, Statistica e Finanza

Dottorato di Ricerca in

Scienze Economiche ed Aziendali

XXXIII Ciclo

THREE ESSAYS OF APPLIED ECONOMICS: SUPERSTAR EFFECT, TAXATION AND WORKERS' MOBILITY, PENALTIES AND DETERRENCE

Settore Scientifico Disciplinare SECS-P/01 ECONOMIA POLITICA

Coordinatore: Ch.mo Prof. Franco Rubino

Supervisore/Tutor: Ch.mo Prof. Vincenzo Scoppa

Firma oscurata in base alle linee guida del Garante della privacy

Dottorando: Dott. Giacinto Gaetano

Firma oscurata in base alle linee guida del Garante della privacy

Abstract: This thesis is composed of three distinct essays, aiming at studying three different phenomena, highly relevant in the economic literature: the phenomenon of the superstar effect, the relationship between the levels of taxation and the mobility of workers across countries or regions, the deterrence resulting from the tightening of laws in a given legal system. The thesis is made up of three chapters. First of all, collecting data from the American basketball championship (NBA), we are able to show the importance of the phenomenon called "Superstars Effects" in the determination of players' wages. In the second chapter, taking into consideration the United States of America, we determine, through the use of Multinomial Logit Models, how the different levels of taxation in the different US states affect the choice of where to play by the players of the NBA championship.

Finally, the third chapter examines the impact of the introduction of Law no. 41/2016 in Italy (the law introducing the crime of "Vehicular homicide"), on the number of accidents and on the main forms of driving offences involved in the introduction of the law, through the use of a Regression Discontinuity Design.

Abstract: Questa tesi è composta da tre saggi distinti, aventi come obiettivo lo studio di tre differenti fenomeni, ampiamente diffusi all'interno della letteratura economica: la formazione del cosiddetto "effetto superstar", la relazione che intercorre tra i livelli di tassazione e la mobilità dei lavoratori in una determinata area geografica, la deterrenza derivante dall'inasprimento di norme all'interno di un ordinamento giuridico.

La tesi è suddivisa in tre capitoli. In primo luogo, collezionando i dati sul campionato di pallacanestro americano (NBA), siamo abili a mostrare l'importanza del fenomeno definito "Superstars Effects" nella determinazione dei salari dei giocatori. Nel secondo capitolo, prendendo in considerazione gli Stati Uniti d'America, determiniamo, attraverso l'utilizzo di Multinomial Logit Models, come i diversi livelli di tassazione praticati nei differenti stati americani incidono sulla scelta della squadra da parte dei giocatori del campionato NBA.

Infine, il terzo capitolo esamina l'impatto dell'introduzione della legge n. 41/2016 in Italia (la norma che ha definito il reato di "Omicidio Stradale"), sul numero di incidenti stradali e sulle principali forme di violazioni del codice stradale, attraverso l'utilizzo dei modelli RDD.

THREE ESSAYS OF APPLIED ECONOMICS: SUPERSTAR EFFECT, TAXATION AND WORKERS' MOBILITY, PENALTIES AND DETERRENCE

Author: Giacinto Gaetano

Supervisor: Vincenzo Scoppa

UNIVERSITÀ DELLA CALABRIA

TABLE OF CONTENTS

INDEX OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

INTRODUZIONE

La tesi è composta da tre differenti ricerche pertinenti al campo dell'economia applicata, con le quali, attraverso l'utilizzo dei più efficienti strumenti statistici a disposizione, si cerca di definire e di spiegare i comportamenti di tre differenti collettivi in tre distinti scenari.

Il leitmotiv comune dei tre lavori è rappresentato dalla volontà di voler definire quali possano essere i principali driver che spingono ed orientano gli individui a prendere decisioni, utilizzando differenti background per l'implementazione delle analisi.

Prendendo spunto dalle recenti introduzioni in termini di economia cognitiva e comportamentale, risulta utile analizzare i processi che vengono impiegati dagli individui per modellare le loro preferenze e prendere le scelte economiche.

La tesi si compone, dunque, di tre capitoli.

Il primo capitolo (*Superstar Effect for Basketball Players*), analizza il fenomeno economico noto come "Effetto Superstars".

Tale concetto è stato formalizzato per la prima volta dall'economista Sherwin Rosen (1981), allo scopo di motivare come un numero relativamente esiguo di individui consegue enormi guadagni ed assume una posizione dominante nel settore in cui sono impiegati. Attraverso l'analisi di questo fenomeno è possibile dare, quantomeno in parte, una spiegazione alle crescenti disuguaglianze di reddito che si istaurano nei paesi più sviluppati e più specificatamente all'interno di determinati settori economici.

Infatti, queste differenze sono più marcate all'interno di particolari ambienti del mercato del lavoro: l'industria dell'intrattenimento (sport, music, cinema). All'interno di questi settori, gli individui definiti "superstar" percepiscono redditi di gran lunga superiore alla stragrande maggioranza dei loro colleghi considerati di "livello inferiore", probabilmente poco proporzionali rispetto alle specifiche differenze in termini di talento.

Rosen spiega che queste differenze derivano dal fatto che le “superstars” sono “sostituti imperfetti” e che, grazie alle innovazioni tecnologiche, i migliori performer sono in grado di servire un mercato sempre più ampio e quindi di raccogliere una quota maggiore dei ricavi presenti. Conseguentemente, questo fenomeno riduce la fetta di guadagno a disposizione dei “co-workers” meno talentuosi.

In aggiunta, in letteratura sono presenti due differenti teorie che desiderano spiegare il processo che porta un individuo a diventare superstar, elaborate rispettivamente da Rosen (1981) e da Adler (1985). Rosen suggerisce che le differenze di talento, seppur minime, generano grandi differenze in termini di guadagni, mentre Adler afferma che gli individui potrebbero avere un talento pressoché simile, ciò che effettivamente incide è la popolarità sulle differenze in termini di reddito.

Per questo motivo, il primo capitolo cerca di rispondere a due differenti research questions, considerando il campionato di pallacanestro più famoso ed importante al mondo (la NBA). In primo luogo, viene analizzato se effettivamente è possibile osservare il fenomeno dell’effetto superstar nella determinazione dei salari dei giocatori, secondariamente se queste grandi differenze di guadagni sono maggiormente spiegabili attraverso quanto affermato da Rosen (importanza del talento) o da Adler (predominanza della popolarità, misurata attraverso il numero di followers su Twitter).

Sfruttando una grande quantità di dati sugli stipendi e una serie di misurazioni delle prestazioni per 8 stagioni NBA, determiniamo che i top performer - definiti in vari modi - guadagnano uno stipendio più alto e sproporzionato rispetto agli altri “buoni giocatori” e troviamo che sebbene la popolarità abbia un’influenza positiva sui salari, le prestazioni (il talento) risultano essere più importanti.

Il secondo capitolo (*Taxation and Workers’ mobility: Evidence from US Basketball*), studia il fenomeno della mobilità geografica dei lavoratori in risposta ai diversi livelli di tassazione sui redditi. La letteratura economica è ricca di studi nei quali si cerca di carpire se e come le differenze in termini di fiscalità riescano ad influenzare le scelte di localizzazione dei lavoratori. Questo fenomeno migratorio, se ben identificato, potrebbe avere importanti risvolti nella progettazione delle politiche fiscali da parte dei legislatori.

I principali contributi presenti in letteratura, mostrano evidenze empiriche principalmente considerando soggetti “top income”. Schimdheiny (2006), Schimdheiny and Slotwinski, (2018), Martinez (2017) hanno presentato prove empiriche sulla Svizzera, sottolineando che i lavoratori “top earners” sono fortemente “mobili” all'interno dei cantoni.

In diverse occasioni, gli autori hanno sfruttato lo sport come veicolo per l’implementazione delle proprie analisi: Klaven et al. (2013) analizzano l’effetto delle differenti aliquote fiscali sulla migrazione internazionale dei giocatori di calcio in 14 paesi Europei nel periodo di tempo tra il 1985 e il 2008, formalizzando una forte relazione negativa.

Il secondo capitolo, dunque, cerca di studiare le risposte migratorie alle differenti aliquote fiscali presenti all’interno del territorio statunitense, sfruttando come contesto di riferimento le scelte di localizzazione intraprese dai giocatori della NBA nel periodo di tempo compreso tra le stagioni 1995/1996 e 2011/2012. È d’uopo sottolineare che gli Stati Uniti rappresentano un'importante opportunità per lo studio di questi meccanismi, in quanto sono caratterizzati dalla presenza di un sistema di tassazione che permette agli stati federali di applicare una propria aliquota fiscale sui redditi.

A tal proposito, ad esempio, Moretti and Wilson (2017) hanno quantificato quanto sensibile fosse la migrazione effettuata da parte dei migliori scienziati americani alla luce delle differenze in termini di tasse sui redditi delle persone fisiche e tasse sui redditi delle imprese, tra gli stati americani.

Nello specifico, il secondo capitolo si propone di analizzare due differenti obiettivi: in primo luogo, se le differenze nella tassazione sui redditi influenzano le scelte di giocare per una data squadra (in un dato stato) da parte dei giocatori della National Basketball Association e, in secondo luogo, se le squadre, eventualmente, fossero “costrette” ad adeguare la loro retribuzione per compensare le specifiche differenze in termini di tassazione.

Per raggiungere questi obiettivi, abbiamo utilizzato diversi modelli statistici, tra i quali il modello multinomiale a scelta discreta. I risultati ottenuti mostrano che le differenze fiscali non incidono né sulla determinazione dei salari né sulle scelte dei giocatori.

In conclusione, ciò che risulta essere alla base delle scelte dei giocatori sono le ambizioni personali e le caratteristiche socio-economiche delle città.

Il terzo capitolo (*Increased penalties for causing road accidents and driving offences: evidence from Italy*) affronta un tema ampiamente dibattuto dall'opinione pubblica italiana nell'ultimo decennio: la sicurezza stradale e l'introduzione della legge n.41/2016 che ha istituito il reato di "Omicidio Stradale", promulgata dal Presidente della Repubblica il 23 marzo 2016.

Con questa nuova disciplina, il legislatore ha cercato di inasprire le sanzioni nei confronti di coloro che, contravvenendo alle norme previste dall'ordinamento vigente, provochino, colposamente, la morte di altri. In particolare, la nuova struttura normativa determina pene aggravate qualora il comportamento del conducente sia stato indotto da un'alterazione psicofisica dovuta al consumo eccessivo di alcol e/o stupefacenti, o da una guida pericolosa.

Inoltre, a seguito di omicidio stradale aggravato è stato introdotto l'istituto della revoca della patente di guida per trent'anni ("ergastolo della patente").

Allo scopo di responsabilizzare gli automobilisti, di favorire una migliore convivenza civile lungo le strade del paese e soprattutto di diminuire il numero dei morti e dei feriti, il governo Renzi ha avvertito l'esigenza di creare questa norma adeguata a dissuadere un individuo dal compiere atti illeciti alla guida del proprio mezzo di trasporto.

Questo capitolo intende verificare la "forza" di una specifica politica pubblica, nel modificare nella direzione desiderata i comportamenti e le condizioni di una determinata popolazione di destinatari (Martini and Sisti, 2009), apportando un nuovo contributo nella letteratura esistente sulle relazioni tra interventi legislativi e sicurezza stradale, concentrandosi, in particolare, su alcune dinamiche comportamentali che si possono instaurare tra gli automobilisti, a seguito di un inasprimento delle sanzioni nelle ipotesi di comportamenti illeciti.

L'obiettivo principale risiede nel verificare se la legge utilizzata dal governo italiano, come deterrente per ottenere un minor numero di incidenti e reati stradali, abbia avuto gli effetti desiderati. Per raggiungere questo obiettivo, vengono proposti modelli econometrici utilizzando la tecnica dei quasi-esperimenti, attraverso l'utilizzo della Regression Discontinuity Design (RDD).

I risultati ottenuti mostrano come l'introduzione della nuova normativa non ha influenzato in modo significativo il numero totale di incidenti o il numero di incidenti gravi (morti e feriti), contrariamente a quanto individuato per il numero di persone sanzionate in stato d'ebbrezza: nello specifico, viene identificata l'esistenza di una relazione negativa e statisticamente significativa tra l'introduzione della nuova norma ed il numero di persone sanzionate per un tasso alcolico ematico superiore a quello consentito dall'ordinamento vigente. Nel dettaglio, l'introduzione della legge ha portato, in media, ad una riduzione di 2.628 persone sanzionate giornalmente (- 5.73%).

Alla luce del fatto che le tre ricerche trattano argomenti indipendenti tra loro, l'analisi della letteratura viene mostrata singolarmente e dettagliatamente all'interno di ogni capitolo. La tesi è organizzata come segue. Nel Capitolo 1 viene presentata la ricerca sul concetto di "Effetto Superstar". Nel Capitolo 2 viene esposta l'analisi circa la relazione tra livelli di tassazione e mobilità dei lavoratori. Nel capitolo 3 viene mostrata l'analisi sul presunto effetto deterrente della legge italiana che ha introdotto il reato di "omicidio stradale".

INTRODUCTION

The thesis is composed of three different researches relating to the field of applied economics, with which, through the use of the most efficient statistical tools available, we try to define and explain the behavior of three different collectives in three distinct scenarios.

The common leitmotif of the three works is represented by the desire to define what may be the main drivers that push and guide individuals to make decisions, using different backgrounds for the implementation of the analyzes.

Taking a cue from recent introductions in terms of cognitive and behavioral economics, it is useful to study the processes that are used by individuals to model their preferences and make economic choices.

This work consists of three chapters.

The first chapter (*Superstar Effect for Basketball Players*) considers the economic phenomenon known as “Superstars Effect”.

This concept was first formalized by the economist Sherwin Rosen (1981), in order to motivate the fact that a relatively small number of individuals report huge earnings and assume a dominant position in the sector in which they are employed. Through the analysis of this phenomenon it is possible to give, at least in part, an explanation for the growing income inequalities that arise in more developed countries and more specifically within certain economic sectors.

In fact, these differences are more marked within particular environments of the labor market: the entertainment industry (sport, music, cinema). Within these sectors, individuals defined as "superstars" earn far greater incomes than the vast majority of their colleagues considered to be of "lower level", probably not very proportional to the specific differences in terms of talent.

Rosen explains that these differences are due to the fact that "superstars" are imperfect substitutes and that, thanks to technological changes, the best performers are able to serve a larger market and therefore collect a greater share of revenues. Consequently, this phenomenon reduces the slice of income available to the less talented.

In addition, in the literature there are two different theories that try to explain the process that leads an individual to become a superstar, elaborated respectively by Rosen (1981) and Adler (1985). Rosen suggests that differences in talent, even if small, generate large differences in terms of earnings, while Adler says that individuals may have almost similar talent, what actually affects incomes it is popularity.

For this reason, the first chapter tries to answer two different research questions, considering the most famous and important basketball championship in the world (the NBA).

Firstly, it analysed whether it is actually possible to observe the phenomenon of the superstar effect in determining the wages of players, secondly if these large differences in earnings are more explainable by what stated by Rosen (importance of talent) or by Adler (predominance of popularity, measured by the number of followers on Twitter).

By leveraging a large amount of salary data and a series of performance metrics for 8 NBA seasons, we determine that top performers - defined in various ways - earn a disproportionate higher salary than good players and we find that although popularity has a positive influence on wages, performance (talent) turns out to be more important.

The second chapter (*Taxation and Worker's mobility: Evidence from US Basketball*), studies the phenomenon of the geographical mobility of workers in response to the different levels of income taxation. The economic literature is full of studies in which authors try to understand if and how the differences in terms of taxation are able to influence the location choices of workers. This migration phenomenon, if well identified, could have important implications for the planning of fiscal policies by legislators.

The main contributions in the literature show empirical evidence mainly considering "top income" subjects. Schimdheiny (2006), Schimdheiny and Slotwinski, (2018), Martinez (2017) presented empirical evidence regarding Switzerland, emphasizing that top earners are strongly mobile within cantons.

On several occasions, the authors have used sport as a vehicle for the implementation of their analyzes: Klaven et al. (2013) investigate the effect of different tax rates on the international migration of soccer players in 14 European countries in the period between 1985 and 2008, formalizing a strong negative relationship.

The second chapter, therefore, tries to study the migratory responses to the different tax rates present within the US territory, using as a reference context the localization choices made by the NBA players in the period of time between the 1995/1996 and 2011/2012. It should be emphasized that the United States represents an important opportunity for the study of these mechanisms, as they are characterized by the presence of a taxation system that allows federal states to apply their own tax rate on income.

Moretti and Wilson (2017) quantified how sensitive was the migration carried out by the best American scientists compared to the differences in terms of personal income taxes and corporate income taxes, between American states.

Specifically, the second chapter proposes to examine two different objectives: first, whether the differences in income taxation influence the choices to play for a given team (in a given state) by the players of the National Basketball Association and, second, if the teams, eventually, were "forced" to adjust their pay to compensate for specific differences in terms of taxation.

To achieve these objectives, we used several statistical models, including the multinomial discrete choice model. The results obtained show that the tax differences do not affect either the determination of wages or the choices of the players.

In conclusion, what appears to be at the basis of the choices of the players are the personal ambitions and the socio-economic characteristics of the cities.

The third chapter (*Increased penalties for causing road accidents and driving offences: evidence from Italy*) deals with a topic widely debated by the Italian public opinion in the last decade: road safety and the introduction of the Vehicular Homicide Law (VHL) promulgated by the President on 23 March 2016.

With this new discipline, the legislator has tried to tighten the sanctions against those who, in contravention of the rules laid down by the current legislation, culpably cause the death of others. In particular, the new regulatory structure determines aggravated penalties if the driver's behaviour has been induced by a psychophysical alteration due to excessive consumption of alcohol and / or drugs, or by dangerous driving.

Furthermore, following an aggravated Vehicular Homicide, this law introduces the institution of the revocation of the driving license for thirty years ("life ban of the license").

In order to empower motorists, to favor a better civil coexistence along the country roads and above all to reduce the number of dead and injured, the Renzi government has felt the necessity to create this adequate rule to dissuade an individual from carrying out illegal acts.

This chapter intends to verify the "strength" of a specific public policy in modifying the behaviors and conditions of a given target population in the desired direction (Martini and Sisti, 2009), making a new contribution in the existing literature on relations between legislative interventions and road safety, focusing, in particular, on some behavioral dynamics that can be established in motorists, following a tightening of sanctions in the event of unlawful behavior.

The main objective lies in verifying whether the law used by the Italian government, as a deterrent to obtain fewer traffic accidents and crimes, has had the desired effects. To achieve this goal, we propose econometric models using the technique of quasi-experiments, through the use of Regression Discontinuity Design (RDD).

The results obtained show that the introduction of the new legislation did not significantly influence the total number of accidents or the number of serious accidents (deaths and injuries), contrary to what was identified for the number of people sanctioned in a state of intoxication: specifically, the existence of a negative and statistically significant relationship is identified between the introduction of the new law and the number of people sanctioned for a blood alcohol level higher than that allowed by the current legislation. In detail, the introduction of the law led, on average, to a reduction of 2.628 people sanctioned daily (- 5.73%).

In light of the fact that the three studies deal with independent topics, the analysis of the literature is expressed individually and in detail within each chapter. The thesis is organized as follows. In Chapter 1, research on the concept of the “Superstar Effect” is shown. Chapter 2 presents the analysis of the relationship between taxation levels and worker mobility. Chapter 3 illustrates the study on the alleged deterrent effect of the Vehicular Homicide Law" (VHL).

CHAPTER ONE

SUPERSTAR EFFECTS FOR BASKETBALL PLAYERS

This paper analyzes if superstar effects emerge in the determination of earnings of US basketball players (NBA), that is, we study if the relationship between earnings and performance is convex. In contrast to the existing literature, we exploit a wealth of data on salaries (accurate data deriving from salary caps) and a host of measures of performance (Points, Assists, Rebounds, Blocks, Steals, Turnovers, Personal Fouls, Game Played, Free Throws Percentage, 2-Points Percentage, 3-Points Percentage) for 8 NBA seasons. We find that top-performers – defined in various ways – earn a disproportional higher salary with respect to good players. We also test if top-salaries are related to popularity and we find that although popularity has a positive influence, performance trumps popularity.

JEL Classification Codes: J31; D31; J24; Z20; Z22.

Keywords: Superstar effects; Wage determination; Wage distribution; Sports Economics;

1.1 Introduction

Income inequality has soared in developed countries in recent decades. The share of total income going to the top decile, to the top 1% or even to the top 0.1% has increased dramatically, especially because of growing inequality in the labor income rather than capital income (Atkinson et al., 2011). In the United States, for example, between 1975 and 2012 around 47% of total growth in pre-tax incomes went to the top 1%. The share was also high in a number of other (mostly) English-speaking countries: 37% in Canada and over 20% in Australia and the United Kingdom (Keeley, 2015). The literature for years has debated the relationship between income inequality and growth, furthermore whether redistribution policies, which aim to provide greater equality on disposable incomes, have negative effects on growth. (Cigliano, 2014).

High yields typically derive from a low level of competition and can damage the possibility of having equal opportunities within an economic context (Franzini et al., 2016).

These inequalities are highlighted in a specific sector of labor market: the entertainment industry (sports, pop music, cinema). The key to the high earnings of few people (superstars) comes from the huge audience that they are able to reach due to scale economies (Franck and Nuesch, 2012). The various technological innovations have made it possible to make the performances of some stars available globally, maintaining an objectively low level of cost. Marshall (1947) was the first to point out that technology would make it possible to bring down prices on quality goods that would in turn become predominant in the market.

The economics of superstars was proposed by Sherwin Rosen to explain why "relatively small numbers of people earn enormous amounts of money and seem to dominate the fields in which they engage" (Rosen 1981, 845). Rosen states that superstars are imperfect substitutes and assumes – thanks to technological changes – that the best performers are able to serve a bigger market and thus reap a greater share of its revenue. However, this inevitably reduces the spoils available to the less gifted in the field.

Table 1.1: The Forbes Top 20 Highest-Earning Celebrities

#	Name	Earnings	Category	#	Name	Earnings	Category
1	Floyd Mayweather	$285 mil	Athlete	11	Bruno Mars	$100 mil	Musician
2	George Clooney	$239 mil	Actor	12	Conor McGregor	$99 mil	Athletes
3	Kylie Jenner	$166 mil	Personality	13	Neymar	$90 mil	Athletes
4	Judy Sheindlin	$147 mil	Personality	14	Howard Stern	$90 mil	Personalities
5	Dwayne Johnson	$124 mil	Actor	15	Ellen DeGeneres	$87.5 mil	Personalities
6	U2	$118 mil	Musician	16	James Patterson	$86 mil	Authors
7	Coldplay	$115 mil	Musician	17	LeBron James	$85.5 mil	Athletes
8	Lionel Messi	$111 mil	Athlete	18	Rush Limbaugh	$84.5 mil	Personalities
9	Ed Sheeran	$110 mil	Musician	19	Katy Perry	$83 mil	Musicians
10	Cristiano Ronaldo	$108 mil	Athlete	20	Robert Downey Jr.	$81 mil	Actors

In the literature, there are basically two competing – but not mutually exclusive – theories of superstar formation proposed by Rosen (1981) and Adler (1985). Rosen explains how small differences in talent translate into large differences in earnings, while Adler argues that superstars might even emerge among equally talented performers due to the positive network externalities of popularity, this is because the acquisition of knowledge by consumers imply discussions among consumers themselves and this is favoured by common prior knowledge and familiarity.

Sport has been used in different circumstances in order to obtain empirical evidence in the economic literature, as it often gives the possibility of obtaining data and information, more complex to reach in different sectors. In our case, the NBA represents a perfect system for obtaining significant results, suited to our research and to the reference hypotheses defined by Rosen. First, in the American basketball there are groups of players "superstars", universally more recognized than their colleagues (LeBron James, Stephen Curry, Kevin Durant), who receive salaries much higher than the average. Secondly, the league's popularity has reached unapproachable heights for any other form of sport in the world: the NBA has established itself as one of the reference game not only in the United States but also in Europe and Asia (there are several commercial agreements signed by the association with China), trying powerfully to overcome the barriers, especially technological ones, with the African continent. This great international attention leads to broadcast the games in over 200 different countries and to sell over 15% of the merchandising abroad. Moreover, with a great attention to the internal balance and competitiveness of the league by exploiting technological progress, the National Basketball Association has become a real example to follow in terms of business. The focus on product quality, the relationships with its customers (fans) and the spectacular ability to take positions in difficult situations have allowed the NBA to become an iconic brand and a real business model.

In this paper we have two different research target. Firstly, we analyze if superstar effects emerge in the determination of salaries of basketball players from NBA, that is, we study if the relationship between salary and performance is convex, that is, if salaries increase at increasing rate for star players, reporting different estimates with different proxy measures of superstars. Secondly, we investigate also the alternative explanation of super-earnings based on players' popularity (Adler, 1985), through data on twitter followers. The unique strength of this paper lies in its wide-ranging and accurate dataset: we exploited a very large (eight seasons) and precise dataset (the sources from which we collected the data are official), take advantage of all the main performance and biographical characteristics of the players. In addition, the methodology used in this study is robust compared to previous works in the literature. The paper is organized in the following way. Section 2 discusses the economic literature about the topics. In Section 3 we describe the dataset and present some descriptive statistics. In Section 4 we carry out the empirical analysis concerning performance and Superstar Effects and the relationship between talent and popularity and their impact on salary. Section 5 concludes.

1.2 Literature

The term 'superstars effect' was coined for by Sherwin Rosen in his seminal paper "The Economics of Superstars" (1981). Rosen's concept started from the observation that within certain sectors (television, cinema, team sports, music) there is a concentration of earnings among a few individuals, marked skewness in the associated distributions of income and very large rewards at the top. This phenomenon is realized when a small fraction of individuals earn an enormous amount of money compared to the remaining workers, dominating the activities in which they are involved. There are two common elements in this kind of sectors, where this phenomenon can be observed: first, a close connection between personal reward and the size of one's own market; second, a strong tendency for both market size and reward to be skewed toward the most talented people in the activity (Rosen, 1981). The phenomenon is present in sectors with a massive presence of technology: under certain circumstances, the high technological development has allowed the perfect duplication of products at costs equal to 0 (Marshall, 1947; Rosen, 1981; Adler, 1985), allowing the creation of scale economies that facilitate such high potential earnings (Lucifora & Simmons, 2003).

Starting from Rosen’s paper, Adler (1985) explains that wage differences may also be present in sectors where there are no real differences in talent but exist in sectors where consumption requires knowledge (Adler, 1985). The transfer of knowledge among consumers involves discussion with other consumers, and a discussion become easier if all participants share common prior knowledge. Following Adler, the grounds behind the differences in income are the differences among stars of their level of popularity, admitting that pure luck could be one possible mechanism by which consumers initially choose a particular artist, whereas Rosen did not explain the mechanism by which stars emerge. In addition, Adler (2006) determined that to become a superstar, objectively recognized in a given market environment, is very important the role provided by the publicity and the amount of public appearances, taking into strong consideration the success achieved on social networks.

Over the years, taking inspiration from the studies of Rosen and Adler, other authors have provided their contribution on the topic in the literature. MacDonald (1988) describes the path taken by young artists, whose uncertainty about talent is high, assuming that the gains are a convex and increasing function of a person's talent taking into consideration that this function is stochastic. Borghans & Groot (1998) offer a vision about the elements that bring a superstar to emerge, considering two factors: first, superstars must be more talented than other people; second, the creation of a certain degree of monopoly power is needed which emerges due to the position of number one of the superstars. Being the best develops a monopoly power, which explains the huge salaries of the superstars. The greater compensation is not due to the way the superstar behaves, but to the fact that people are not happy to look at other artists once the superstar has been identified: superstars get a "property right" on the number one position that they occupy. All theories agree that although quality and talent are often difficult to measure and identify, superstars have personal characteristics that are unique.

Several past studies have shown that superstar effects are most likely to be noted in the entertainment business and in professional sports (Adler, 1985; Fort & Quirk, 1995). An important study that tried to test the theories of Rosen and Adler in professional team sports was made by Lucifora and Simmons (2003), with which examined the dynamics that determine wages among soccer players in the Italian league. Given the popularity of some players, called top players, the relationship between individual productivity and retribution can lead to 'superstars' effects. In this context, the marginal revenue produced by a player may be tied to the 'additional' price that a viewer is willing to pay to watch it play (live or on television). In this paper, it is observed the presence of the “superstars effect” using different OLS and FGLS models and data relating to just one season (1995/1996), building the superstar variables through the strike rate (the ratio between goals and appearances in the league). Following the work of Lucifora and Simmons, several authors have studied whether small variations in skills translate into large differences in wages in the best European football leagues, such as the German Bundesliga (Lehmann & Schulze, 2007) and the Spanish Primera Division (Garcia del Barrio & Pujol, 2005). In American team sports, it is possible to have even more detailed information for the achievement of more complete analysis of the “superstar effect”. Given the structure of the different leagues (MLS, NBA, NFL, NHL), the league-designated player rule and all-star game participation provide unique measures of superstar status unattainable in existing studies (Kuethe & Motamed, 2010). There are several implications that can be identified in cases where it is possible to associate superstar status with particular player. Having players recognized as "superstars" in the team allows presidents to earn extra revenues (media, sponsors, merchandising, tickets). This could lead them to push managers to buy players not in relation to their ability to win matches but to their global knowledge (Hausman & Leonard, 1997).

Furthermore, few studies in literature have tried to provide empirical evidence about the relationship between talent and popularity in wage determinations for some categories of individuals, reporting mixed results. Lehman and Schulze (2008), using three measures of performance and the number of citations in the online version of the soccer magazine Kicker (as proxy of popularity), did not find any effects on salaries of players in the German soccer league. On the other hand, Franck and Nuesch (2012), exploiting a dataset with 20 different performance indicators considering data from German football league for five season, found that both talent and popularity increase the demand for star players. Moreover, Carrieri et al. (2018) using different methods of estimations found that performance, popularity and bargaining power are correlated with higher salaries; in particular, popularity overcomes all the others factors.

1.3 Data and Descriptive Statistics

The American National Basketball Association (NBA) was founded in 1946 by the owners of the largest American arenas, with the name of BAA (Basketball Association of American). In 1949 through the union of BAA and other minor leagues, the NBA was born with 17 participating teams. Over the years there were several relocations and changes that led to the creation of a championship with 30 participating teams (29 American and one Canadian), divided into two distinct conferences (eastern and western), according to a geographical criterion, each of which is made up of 3 divisions and each division is composed of 5 teams. Since 2004 during the season there are three different phases: Regular Season, Play-Off and Finals. The NBA Finals represent the last act of each season, decisive for decreeing the winning team of Larry O'Brien Trophy. Annually, like the other American sports characterized by the presence of the franchise-teams, also in the NBA in the month of June the draft lottery is carried out. It is an event that allows each team to choose, following certain criteria, new players that are at least 18 years old, coming from colleges or other international organizations to be included in their roster. The American sports system fixes that the training of young athletes takes place directly in high schools and colleges, excluding the existence of "youth teams", determining that the only great possibility that teams have to 'pick' the young players with the best qualities is through this lottery. The first 14 choices belong to the teams that in the previous season did not have the strength to qualify from the regular season to the playoffs, with an order of choice determined through a lottery that follows certain rules. All players selected for the Draft are considered Rookies (players with the first year of experience in the NBA) and for selected in the first round there is a specially dedicated bargaining system: "the Rookie Scale Contract".

Considering the economic ratings provided by Forbes, the league teams generated $ 7.4 billion in revenue in 2017, up 25% from 2016. The average value of NBA teams is $ 1.65 billion. The New York Knicks is the franchise with the highest rating: $ 3.6 billion. The increase of revenues can be explained by the great internationalization of the league, especially in Europe and China. The main sources of income for teams in the league are: broadcast rights, advertising, merchandising and tickets sales. In 2015, the NBA has signed an agreement with Tencent for the transmission of games and other multimedia contents in China, for a total value of $ 800 million. Since 2016, the contract signed by the league with Turner Broadcasting Inc. and The Walt Disney has been in force, for an economic value of $ 24 billion for nine years. After the new partnership with Nike (an eight-year contract of $1 billion), for the first time the league authorized teams to add a sponsorship logo on player jersey to generate additional revenue. In 2017 the league released an official announcement in which stated that NBA broke attendance record for third straight season with a sellout record of 723 matches in the 2016/2017 regular season, with an increase in revenues from ticket sales. In NBA there is the Revenue Sharing System: all teams collect their annual revenue which will be redistributed from the teams that earned more to the team with lower revenues, with the constraint that can participate only teams that have generated revenue equal to at least 70% of the league average. To demonstrate the popularity of the NBA in the world it is very interesting to observe the data about the number of followers on social networks of the main players. LeBron James is the most-followed American athlete on social media with over Twitter (41 million) and Instagram (36 million).

In our empirical analysis we use data concerning salaries, physical and biographical characteristics of players, and a host of measures of their performance. We acquire data relating to NBA players starting from the 2010/2011 season up to the 2017/2018 season, a full 8 seasons dataset made up of over 3,000 observations. A key element for the access to the data is provided by the rigid regulation of the NBA. In fact, all franchises are obliged, due to the salary cap, to publish and systematically update the salary data of each basket player regularly under contract. This represents one of the most complex systems of management of monetary resources that exist in the world of sport, defined in the NBA Collective Bargaining Agreement. It represents the limit to the total amount of money that the teams belonging to the league are allowed to pay to 'sign' their players. It was established in 1946/47 but for about 40 years it was no used, leaving the team free to spend. It was reinstated in 1984/85 to try to balance the forces and ensure a greater uniformity of income among all the teams. Since 1984 all teams are obliged, therefore, to publish annually and accurately the data related to the salaries that correspond to their players.

The sources used to obtain information on salaries and the main biographical characteristics were the official website of the National Basketball Association (www.nba.com) and the largest online sports team and player contract resource on internet www.spotrac.com.The main technical statistics of the players have been acquired by the ESPN site (Entertainment & Sports Programming Network), one of the most important American networks which transmits sports 24 hours a day.

For each player, we have collected data on salary[1] and on biographical characteristics: height, weight, age, nationality, race; years of experience, role[2], college attendance, number

[1] We use inflation-adjusted salaries (inflation rate from the Bureau of Labor Statistics), transforming nominal salaries in 2018 dollars.

[2] In basketball, as well as in football, players in the field have different roles and tasks. It is possible to determine four different macro-roles in the basket: the Point Guard; the Guard; the Forward and the Center. The point guard (or "shooting guard") represents the evolution that over time has had the role of the playmaker, the player in charge of creating the attack game, generally the shortest player but also the shrewdest and quickest in dictating game times; the Guard represents a very large group of players. In modern basketball this 'spot' is generally intended for the basketball player with better defensive skills; the Forward (we treat indistinctly small forward and large forward) represents the most complete players from a physical point of view, able to be effective both near the basket that in mid-range positions or beyond line of 3 points; the Center (in Europe is often used the expression ‘pivot’) is generally the highest and most robust player of the team, especially called to fight under the basket in the capture of rebounds both in the offensive and defensive midfield

of pick in the draft, participation to the “All Stars Game”. We have also gathered data on the following measures of performance: points, assists, rebounds, steals, blocks, turnovers, minutes for games, games played, percentage of shots goals, percentage of three points goals, percentage of free throws goals, personal fouls, etc. Measures of performance are per- game average. Finally, we gather data on the followers of Twitter for each player (at October 2017). We combine data on performance in year *t* with salaries in year *t+1*. We have available 3,032 observations for 736 players, on average about 4 yearly obs. for each player.

Descriptive statistics are reported in Table 2. The mean salary is 6,553,000, while the median wage is much lower (4,060,000). The age is 27.3. On average players are 201 cm. tall and weigh 101 kg. About 22% are non-US citizens, 74% are black. On average, players have played 55 games per season, have scored 9.4 points per game, 2 assists, 4 rebounds, Have 72% of success in free-throws, 26% in 3-points. 13.2% of them have played in the All-Stars-Game[3]. We define a dummy “Draft: Best Five Picks” equal to one if a player is in the first 5 picks (and 0 otherwise): 15% of players were in the best five. We measure *Popularity* with the number (in thousand) of followers on twitter (on average, 524 thousand).

[3] The NBA All-Star Game is a basketball exhibition game hosted every February, matching a mix of the league's star players. Each team consists of 12 players, making it 24 in total. The starting lineup for each squad is selected by a combination of fan, player, and media voting.

Table 1.2: Descriptive Statistics

	Mean	S.D.	Min	Max	Obs.
Salary	6553.218	6539.771	9.027	37457.15	3,032
Height	201.143	8.994	165.100	238.76	3,032
Weight	101.198	11.807	60.328	139.253	3,032
Age	27.293	4.165	20	41	3,032
Experience	5.020	4.191	0	21	3,028
non-US	0.221	0.415	0	1	3,032
Black	0.738	0.440	0	1	3,032
Forward	0.341	0.474	0	1	3,032
Center	0.209	0.407	0	1	3,032
Guard	0.245	0.430	0	1	3,032
Point Guard	0.204	0.403	0	1	3,032
Points	9.415	5.747	0	32	3,032
Assists	2.038	1.873	0	11.7	3,032
Blocks	0.459	0.461	0	3.7	3,032
Rebounds	3.990	2.485	0	15.7	3,032
Steals	0.712	0.429	0	2.5	3,032
Personal Fouls	1.882	0.696	0	6	3,032
Turnovers	1.291	0.796	0	5.7	3,032
Free Throw Perc.	0.724	0.153	0	0.962	2,991
3-Point Field Goal	0.264	0.155	0	0.8	3,017
Field Goal Perc.	0.449	0.074	0	0.75	3,028
Games Played	55.103	21.353	1	82	3,032
Year	2014.574	2.222	2011	2018	3,032
All-Star Game	0.133	0.340	0	1	3,032
Draft: Best 5 Pick	0.153	0.360	0	1	3,032
Popularity	524.769	2426.693	0	38300	3,032

Source: www.nba.com; www.espn.com/nba/teams; www.spotrac.com

We first show graphically some preliminary evidence on the skewness of the distribution of salary. In Figure 1 we show the long right-tail in the distribution of salaries, that is, the right skewness. The positive asymmetry of the distribution is confirmed by the fact that the median salary (4060) is much smaller than the average salary (6553). The ratio between the 90 percentile and the 10 percentile is 16.4. Figure 2 shows that average salary earned by each decile: it is immediately clear how much higher are the salaries earned by the top deciles with respect to lower deciles.

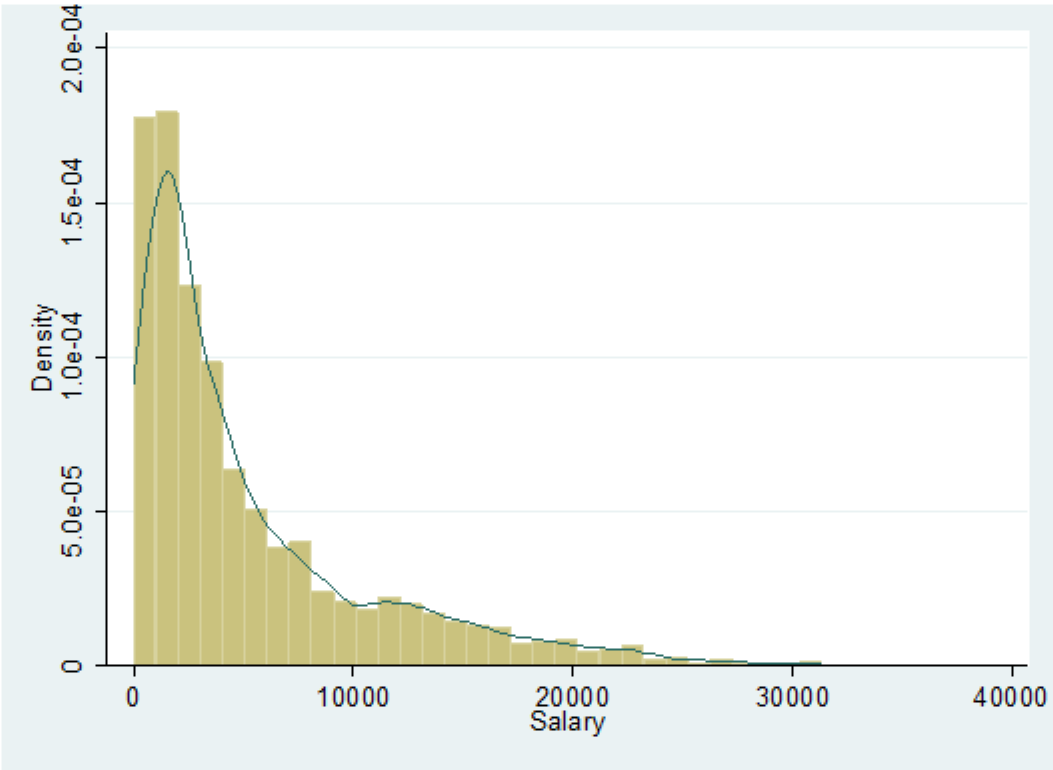

Figure 1.1: The skewness in the distribution of salaries

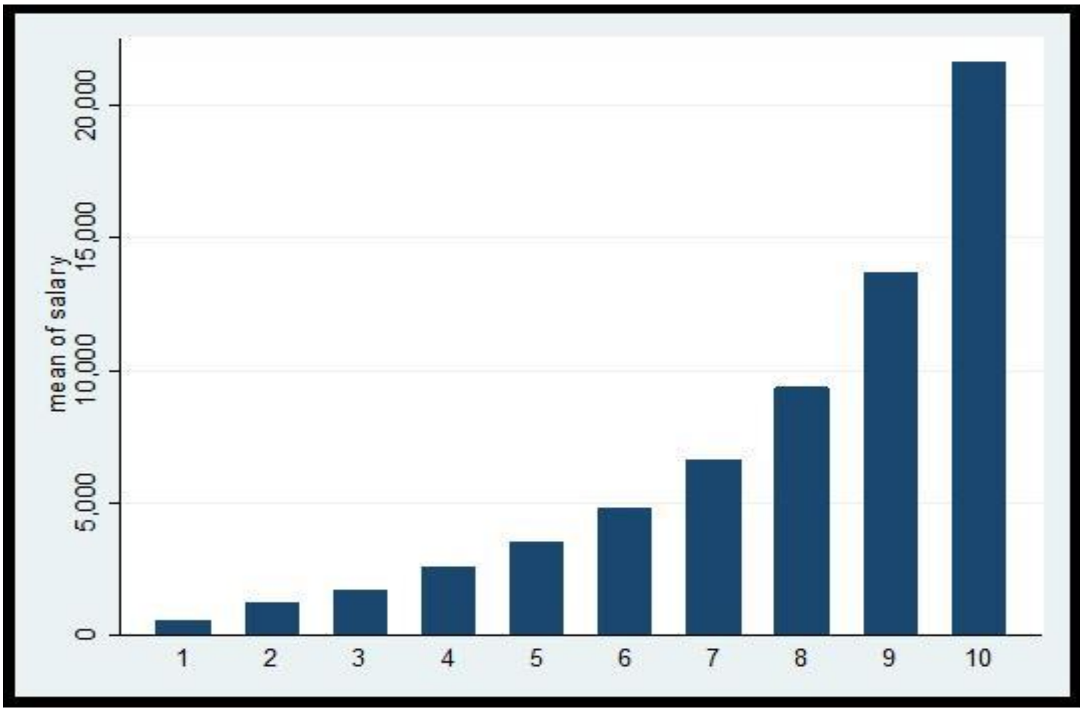

Figure 1.2: Mean of Salary by decile

1.4 Empirical Analysis and Results

1.4.1 Superstar Effects and Performance

In this Section we carry out our empirical analysis to investigate if superstars earn disproportionately more than remaining players. In contrast to the existing literature, we do not use as a dependent variable the *log* of salaries. In fact, if the real relationship is convex, taking the log could kill the convexity. Suppose salary and performance are related through this simple and highly convex relationship: $salary = exp(\beta Performance)$. Then, by taking the log one would obtain: $ln(salary) = \beta Performance$, that is, a linear relationship *between ln(salary)* and performance. On the other hand, if one estimate the typical equation used in the literature: $ln(salary) = \beta_0 + \beta_1 Performance + \beta_2 (Performance)^2$ and is not able to reject the null hypothesis that $\beta_2 = 0$, it is nonetheless evidence of an increasing and convex relationship if $\beta_1 > 0$.[4]

Therefore, we use *Salary* in linear form as a dependent variable:

$$Salary_{it} = \beta_0 + \beta_1 Performance_{it} + \beta_2 (Performance_{it})^2 + X_{it} + \lambda_t + \mu_T + \varepsilon_{it}$$

where *Performance* are, in turn, several measures of performance, *X* is a vector of individual characteristics (age, years of experience, height, weight, race, nationality, etc.),[5] λ are year dummies and μ are team fixed effects. Standard Errors are clustered at player level and robust to the heteroskedacity.

First of all, we use as a measure of performance arguably the most important indicator: the number of points scored. OLS estimates are reported in Table 1.3. In column (1) of Table

[4] Since $salary = exp(\beta_0 + \beta_1 Performance)$, then the first derivative is $\frac{\partial salary}{\partial Performance} = exp(\beta_0 + \beta_1 Performance)\beta_1 > 0$; while the second derivative is: $\frac{\partial^2 salary}{\partial Performance^2} = \exp(\beta_0 + \beta_1 Performance)(\beta_1)^2 > 0$.

[5] Since Height and Weight are highly correlated (ρ=0.81) to avoid multicollinearity problems we use only Height; similarly, we only use Age omitting Years of Experience (ρ=0.88).

3 we simply estimate salaries in relationship to *Points* and *Points Squared*, without any controls. We show that salaries increase with points at an increasing rate, that is, a convex function emerges, since both β_1 (374.3) and β_2 (16.3) are positive and highly significant (*t*-stat=5.59 and *t*-stat=5.60, respectively). From column (2) we control for *Height*, *Age*, *Age squared*, *Black* and *non-US*. We find that height increases salaries (about 1 million more for 10 cm.), while salaries are increasing with age until the age of 31.9, then decrease. Salaries of black players are not statistically different from others, while non-US players earn about 1.1 million more than US players. More importantly, the impact of points on salary is highly non-linear: scoring a point more gives 419,000 dollars more at the 10th percentile of the distribution of points (2.8), 621,000 at the median of points (8.4) and 946,000 at the 90th percentile. In column (3) we also control for the role of players and for year dummies. Our results are almost unchanged. Finally, in column (4) we additionally control for team fixed effects (30 teams) to verify if the differences in players' earnings derive from teams paying different wages. We are not able to reject the null hypothesis that team fixed effects are null (F=0.68 with a p-value=0.89) and the convex impact of points scored on earning remain unchanged.

Table 1.3: The Relationship between Salary and Points. OLS Estimates

	(1)	(2)	(3)	(4)
Points	374.349***	317.747***	317.857***	330.975***
	(67.020)	(62.166)	(61.620)	(63.786)
Points sq.	16.302***	18.067***	17.794***	17.344***
	(2.912)	(2.736)	(2.793)	(2.872)
Height		104.468***	70.930**	71.499**
		(16.886)	(29.424)	(28.786)
Age		2662.776***	2637.020***	2659.739***
		(268.928)	(270.666)	(266.077)
Age sq.		-41.679***	-41.260***	-41.564***
		(4.731)	(4.761)	(4.685)
Black		393.113	328.340	343.067
		(314.555)	(301.012)	(305.127)
non-US		1095.665***	862.062***	913.712***
		(343.323)	(328.572)	(322.600)
Center			949.586**	947.116**
			(408.487)	(400.462)
Guard			-375.381	-345.744
			(407.491)	(405.429)
Point Guard			-236.153	-246.169
			(530.376)	(531.042)
Constant	1046.203***	-61085.755***	-53084.409***	-53671.800***
	(284.357)	(5195.365)	(7545.251)	(7505.745)

Year dummies	NO	NO	YES	YES
Team dummies	NO	NO	NO	YES
Obs.	3032	3032	3032	3032
Adjusted R^2	0.460	0.535	0.556	0.557

Notes: The dependent variable is Salary. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

An alternative way to verify the convexity of the relationship between points and salaries is to build a dummy for each quintile of points and regress salaries on the quintile dummies, leaving the first quintile as the reference category. We estimate the same specifications of Table 3, using the quintile dummies instead of Points and Points Squared. In Table 1.4 we show that the differences between a quintile and the previous are always increasing. Considering column (3), we see that the second quintile earns 877 thousand more than the first quintile, the third quintile earns 1,312 (=2190-877) thousands more than the second quintile, the fourth quintile earns 2,732 thousand more than the third quintile, and finally the fifth quintile earns 6,151 thousand more than the fourth quintile. All these differences are highly statistically significant.

Table 1.4: Quintile Dummies and Salaries. OLS Estimates.

	(1)	(2)	(3)	(4)
2nd quintile pts	1369.173***	1013.049***	877.558 ***	924.502 ***
	(191.410)	(192.350)	(193.588)	(196.923)
3th quintile pts	2741.177***	2307.368***	2190.341***	2256.718***
	(222.556)	(227.405)	(223.830)	(226.921)
4th quintile pts	5437.184***	5012.879***	4922.575***	5049.395***
	(329.422)	(309.684)	(296.933)	(301.429)
5th quintile pts	11668.77***	11246.88***	11073.79***	11122.78***
	(491.957)	(455.271)	(459.515)	(444.977)
Observations	3032	3032	3032	3032
Adjusted R^2	0.400	0.475	0.498	0.501

Notes: The dependent variable is Salary. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Furthermore, one can check if – controlling for Points – there exists a discontinuous jump in the salary of players scoring at the top deciles, using the dummies Star: 91-100 Percentiles and the dummy Star: 81-90 Percentiles. In Table 1.5 we see that Points has a strong positive effect on salaries (about 600 thousand dollars more for each point), but there is a jump in the salary for players scoring in the Decile 81-90 (+1685) and for those scoring in the top decile (+3435). In Table 1.6 we only use the dummy Star: Decile 91-100 and we find again that a discontinuity for top players.

Table 1.5: Top Two Deciles. OLS Estimates.

	(1)	(2)	(3)	(4)
Star: 91-100 Percentiles	3125.179***	3429.488***	3435.213***	3353.152***
	(716.924)	(679.880)	(657.355)	(666.792)
Star: 81-90 Percentiles	1494.385***	1670.702***	1685.345***	1651.982***
	(541.580)	(490.273)	(473.096)	(465.397)
Points	620.254 ***	591.389 ***	584.658 ***	590.099 ***
	(35.118)	(33.156)	(32.296)	(31.640)
Observations	3032	3032	3032	3032
Adjusted R^2	0.456	0.530	0.552	0.553

Notes: The dependent variable is Salary. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.6: Top Decile. OLS Estimates.

	(1)	(2)	(3)	(4)
Star: 91-100 Percentiles	2095.076***	2272.550***	2265.073***	2206.596***
	(615.828)	(578.367)	(568.836)	(575.550)
Points	685.647 ***	665.256 ***	659.118 ***	663.219 ***
	(30.146)	(27.518)	(26.854)	(26.652)
Observations	3032	3032	3032	3032
Adjusted R^2	0.453	0.527	0.548	0.549

Notes: The dependent variable is Salary. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.4.2 Superstar Effects using several measures of performance (Score)

In contrast to the existing literature that uses only one or two measures of performance, our dataset has several measures of performance. Therefore, we can measure performance not only with the number of Points scored, but also with a host of indicators. We use the (per match) average number of Points, Assists, Rebounds, Blocks, Steals, Turnovers (Lost

Balls), Personal Fouls, Game Played, Free Throws Percentage, Field Goal Percentage (2-Points) Percentage, 3-Points Percentage. To avoid collinearity problems, we build a single index of performance (denoted as Score), by undertaking a principal component analysis summarizing the 11 available measures of performance. Principal component analysis creates linear combinations of the original variables which capture the greatest variance. We only use the first principal component. Score is highly correlated to Points (0.87), Assists (0.69), Rebounds (0.67), Steals (0.76), games played (0.68) and so on. We replicate the analyses above using this time Score as a comprehensive measure of performance instead of Points. In Table 1.7 we find again evidence of the convex relationship between Score and Salary.

Table 1.7: Convexity using Score. OLS Estimates

	(1)	(2)	(3)	(4)
Score	1872.993***	1787.692***	1787.512***	1799.771***
	(72.715)	(66.706)	(66.942)	(65.052)
Score Sq.	203.711 ***	235.943 ***	237.726 ***	241.745 ***
	(27.751)	(27.695)	(27.843)	(27.039)
Observations	2973	2973	2973	2973
Adjusted R^2	0.410	0.487	0.513	0.516

Notes: The dependent variable is Salary. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 1.8 we verify that higher quintiles of scores lead to disproportional higher salaries and in Table 1.9 one can note that there are discontinuities in the relationship between Score and Salary for the top two deciles of Score.[6]

[6] We obtain very similar results if instead of building *Score* with 11 variables, we build Score with a reduced set of variables (Points, Assists, Blocks, ...)

Table 1.8: Quintiles of Score. OLS Estimates

	(1)	(2)	(3)	(4)
2nd quintile Score	1264.203***	629.762 ***	469.393 **	479.323 **
	(175.483)	(196.976)	(197.278)	(202.347)
3th quintile Score	3274.652***	2451.541***	2302.919***	2353.090***
	(243.544)	(249.727)	(245.570)	(248.487)
4th quintile Score	5616.614***	5044.328***	4942.500***	4983.765***
	(334.493)	(320.958)	(314.421)	(319.018)
5th quintile Score	11246.39***	10626.84***	10585.27***	10649.95***
	(512.182)	(467.622)	(480.476)	(461.033)
Observations	2973	2973	2973	2973
Adjusted R^2	0.366	0.442	0.471	0.475

Notes: The dependent variable is Salary. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.9: Top Two Deciles of Score. OLS Estimates

	(1)	(2)	(3)	(4)
Star:91-100 Percentiles	4790.306***	5580.646***	5714.707***	5747.519***
	(729.223)	(667.674)	(664.574)	(650.032)
Star: 81-90 Percentiles	2251.919***	2438.128***	2606.354***	2653.754***
	(575.988)	(514.410)	(501.924)	(495.347)
Score	1392.026***	1237.507***	1217.971***	1222.829***
	(81.046)	(78.474)	(77.158)	(77.425)
Observations	2973	2973	2973	2973
Adjusted R^2	0.398	0.472	0.499	0.502

Notes: The dependent variable is Salary. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A further measure of superstar we are able to use is the participation to the All-Star Game. We estimate the usual specifications in Table 10, first controlling for Points (col. 1-4), and then for Score, (col. 5-8), adding the dummy variable All-Star Game. We find that – also controlling for a comprehensive set of measures of performance – the players participating in the All-Star Game are able to earn a much higher salary then similar performing colleagues: about 4-5 million more.

Table 1.10: All-Star Game. OLS Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Points	620.049***	603.633***	594.023***	594.644***				
	(28.498)	(25.530)	(25.307)	(25.666)				
All-Star Game	4358.467***	4345.623***	4460.204***	4496.945***	5545.709***	5669.811***	5862.422***	5920.126***
	(639.404)	(527.390)	(524.862)	(528.604)	(618.901)	(533.897)	(540.772)	(531.684)
Score					1451.560***	1359.545***	1350.154***	1353.572***
					(74.970)	(70.740)	(69.142)	(69.984)
Observations	3032	3032	3032	3032	2973	2973	2973	2973
Adjusted R^2	0.484	0.556	0.578	0.579	0.435	0.501	0.530	0.533

Notes: The dependent variable is Salary. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We obtain similar results if we use as a measure of superstar a dummy variable indicating if a player has been ranked in the **best 5 picks of the draft**: the impact is about 1.7 million dollars, controlling for Points or about 2.4 when we control for Score, our comprehensive measure of performance (estimates not reported).

Table 1.11: Best Draft and Points

	(1)	(2)	(3)	(4)
Points	725.867***	709.684***	702.859***	704.978***
	(27.146)	(25.282)	(25.498)	(24.729)
Draft: Best 5 Picks	1731.181***	1740.661***	1785.421***	1781.533***
	(496.756)	(470.021)	(460.682)	(454.684)
Observations	3032	3032	3032	3032
Adjusted R^2	0.456	0.529	0.551	0.552

Notes: The dependent variable is Salary. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.12: Best Draft and Score

	(1)	(2)	(3)	(4)
	Salary	Salary	Salary	Salary
Score	1796.037***	1705.465***	1702.022***	1712.795***
	(77.377)	(75.554)	(76.385)	(74.908)
Draft: Best 5 Picks	2244.129***	2436.466***	2426.336***	2427.041***
	(538.018)	(537.030)	(539.249)	(521.482)
Observations	2973	2973	2973	2973
Adjusted R^2	0.387	0.454	0.479	0.481

Notes: The dependent variable is Salary. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.4.3 Talent (Rosen) vs. Popularity (Adler)

In Table 1.13 we try to investigate if the top salaries are caused mainly by a superior talent (as Rosen famously argued) or if the cause of high salaries are due to players' popularity. In doing this, we use as a measure of popularity of each player the number of followers on Twitter[7], which we call Popularity. We run the specifications in Table 10, using as a measure of talent the variables Points and All-Star Game. In addition, we use Popularity. We find that having a superior talent determines an increase in the salary of about 4 million dollars (highly statistically significant, t-stat=7,589), controlling for Points. When controlling for Popularity, the magnitude of the effect is only slightly reduced with respect to the estimates in Table 9 (the effect was about 4.4 million). In addition, the variable "popularity" is in turn highly significant, an increase in the number of followers on twitter for basketball players means an increase in the level of wages.

Furthermore, in order to study what is most important (talent or popularity) in determining wages among American basketball players, we exploit the results obtained through the standardized beta coefficients.

The beta coefficients are used to compare the impact of each individual independent variable to the dependent variable. The higher is the value of the beta coefficient, the

[7] Twitter is a social networking and microblogging service on which people post and interact with short messages called "tweets". The latter were originally made of 140 characters. This social network was created in 2006 and in 2013 become one of the ten most visited websites of world. Nowadays, Twitter has more than 321 million monthly active users and it is very popular among celebrities.

stronger is the effect. These coefficients are all measured in standard deviations, instead of the units of the variables, in such a way that it could be comparable to one another. To simplify, the beta coefficients are the coefficients that you would obtain if the outcome and predictor variables were all transformed in standard scores, also called z-scores, before running the regression. The coefficients are calculated by subtracting the mean from the variable and dividing by its standard deviation.

In Table 1.13, a one standard deviation increase in popularity would yield a 0.077 standard deviation increase in the predicted salary, on the other hand a one standard deviation increase of points and all-star game (proxies of talent) determine respectively an increase in the predicted salary of 0.503 and 0.211.

Therefore, we found that both talent and popularity have a positive and statistically significant relationship with salary, but more specifically the contribution provided by the talent variable is much higher than the contribution deriving from the different levels of popularity.

Table 1.13: Talent vs. Popularity. OLS Estimates

	(1)	(2)	(3)	(4)
Points	598.485***	584.177***	572.711***	572.739***
	(28.925)	(26.573)	(26.080)	(26.515)
All-Star Game	3924.463***	3933.480***	4055.425***	4105.459***
	(653.790)	(530.488)	(534.380)	(535.144)
Popularity	0.221***	0.197***	0.206***	0.214***
	(0.061)	(0.062)	(0.064)	(0.066)
Observations	3032	3032	3032	3032
Adjusted R^2	0.489	0.556	0.582	0.584
Standardized beta coefficients				
	(1)	(2)	(3)	(4)
Points	0.526***	0.513***	0.503***	0.503***
	(28.925)	(26.573)	(26.080)	(26.515)
All-Star Game	0.204***	0.204***	0.211***	0.213***
	(653.790)	(530.488)	(534.380)	(535.144)
Popularity	0.082***	0.073***	0.077***	0.079***
	(0.061)	(0.062)	(0.064)	(0.066)
Observations	3032	3032	3032	3032
Adjusted R^2	0.489	0.556	0.582	0.584

Notes: The dependent variable is Salary. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.5 Conclusion and Further Research

The results obtained and the methodology used give this paper significant standing within literature about the phenomena of wage determination, in particular within labour markets characterised by the presence of large gains and great international visibility. The choice to exploit this accurate and large dataset has allowed us to achieve robust results compared to those already present in the literature. In general, we found evidence regarding the main variables that contribute to the determination of wages for NBA athletes, confirming in some circumstances the results present in several previous papers.

In this paper we developed two different research target. Firstly, we analyze if superstar effects emerge in the determination of salaries of basketball players from NBA, that is, we study if the relationship between salary and performance is convex, that is, if salaries increase at increasing rate for star players, reporting different estimates with different measures of superstars. Secondly, we investigate also the alternative explanation of super-earnings based on players' popularity (Adler, 1985), through data on twitter followers. Regarding the first research question, this paper showed how the superstar effects, according to the lines proposed by Rosen (1981), can be estimated empirically using different models with different proxies' variables of the players' talent. We considered as measure of wage in our models the basic salary of each player, from which, by NBA regulation, additional components related to performance are excluded, such as bonuses for the achievement of certain results that are dominant in the players' contracts in the world of football (literature in the past used almost exclusively football as a vehicle to capture the superstar effect defined by Rosen). The proxies variables of talent that we used are: points, a synthetic indicator formed by 11 different performance measures called score, participation to the All-Star Game, a dummy variable indicating if a player has been ranked in the best 5 picks of the draft.

In all these regressions we have identified positive and statistically significant relationships between the above mentioned variables and the annual remuneration of NBA players.

In addition, in order to more concretely verify the existence of a possible discontinuity (jump) in the salary levels between the players belonging to the richest band and the remaining part, for the variables points and score we implement a further analysis through the division of the sample into quintiles and percentiles, bringing to light in both models a clear difference in remuneration between the top 10% and the remaining players, using a very robust set of control variables, ascertaining again the existence of the superstar effect defined by Rosen based on different levels of talent and confirmed by previous works in the literature.

The next important step of this paper was developed in the second research question. First, we tested the definition of superstar effect provided by Adler (1981), namely that the differences in remuneration are realized due to the presence of positive externalities (popularity) taking into account that players could possess similar or equal talent values. Using the strength and development of Twitter, observing the number of followers present in personal profiles allowed us to directly determine a measure of the global popularity of the different players in the league. The results obtained clearly show the existence of a positive relationship between the variable popularity and the dependent variable of our model.

The last step of the work consisted in determining comparing the different channels affecting the superstar effects. Using standardized beta coefficients, we showed how talent, expressed in points and participation to the all-star game, turns out to be much more influential on salary than the different levels of popularity.

This work could be a starting point for subsequent implementations. It would be very interesting to try to understand the benefits and the economic motivations that push the different general managers sometimes to give higher salary levels compared to the performances actually expressed by the players inside the playing field, what could be the economic drivers directly related to the presence of a specific player in the roster (television audience, level of attendance, contracts of sponsorship).

Furthermore, bearing in mind the starting hypotheses provided by Rosen and Adler, concerning the possible presence of the superstar effect in determining the wages of a certain category of workers, due to the growing inequalities, it could be very useful to try to extend this type of analysis also in other economic sectors (not necessarily super-rich) that have similar characteristics to those prescribed by Rosen and Adler.

1.6 References

Adler, M. "Stardom and Talent." American Economic Review, 75(1), 1985, 208–12.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. "Top Incomes in the Long Run of History." Journal of Economic Literature, 49 (1): 3-71.

Bryson, A., Rossi, G., & Simmons, R. (2014). "The migrant wage premium in professional football: a superstar effect?". Kyklos, 67(1), 12-28.

Carrieri, V., Principe, F., Raitano, M., (2018), "What makes you 'super-rich'? New evidence from an analysis of football players' wages", *Oxford Economic Papers*, 70, 4, pp. 950–973.

Cingano, F. (2014), "Trends in Income Inequality and its Impact on Economic Growth", OECD Social, Employment and Migration Working Papers, No. 163, OECD Publishing.

Coates, D., Frick, B., & Jewell, T. (2016). "Superstar salaries and soccer success: The impact of designated players in major league soccer". Journal of Sports Economics, 17(7), 716-735.

Deutscher, C. and Buschemann, A. (2016) "Does performance consistency pay off financially for players? Evidence from the Bundesliga". Journal of Sports Economics, 17, 27–43.

Franck, E., & Nüesch, S. (2012). "Talent and/or popularity: what does it take to be a superstar?". Economic Inquiry, 50(1), 202-216.

Hausman, J. and Leonard, G., 1997, "Superstars in the National Basketball Association: Economic Value and Policy," Journal of Labor Economics 15, no. 4, pp. 586-624.

Keeley, Brian (2015), "What's happening to income inequality?", in Income Inequality: The Gap between Rich and Poor, OECD Publishing, Paris.

Krueger, A. B. (2005). "The economics of real superstars: The market for rock concerts in the material world". Journal of Labor Economics, 23(1), 1-30.

Kuethe, T. H., & Motamed, M. (2010). "Returns to stardom: evidence from US major league soccer". Journal of Sports Economics, 11(5), 567-579.

Lehmann, E. E., & Schulze, G. G. (2008). “What does it take to be a star? The role of performance and the media for German soccer players”. Applied Economics Quarterly, 54(1), 59-70.

Lucifora, C., & Simmons, R. (2003). “Superstar effects in sport: Evidence from Italian soccer”. Journal of Sports Economics, 4(1), 35-55.

Rosen, S. (1981). The economics of superstars. The American Economic Review, 71, 845-858.

CHAPTER TWO

TAXATION AND WORKERS' MOBILITY: EVIDENCE FROM US BASKETBALL PLAYERS

Taxes influence many economic decisions such as consumption, savings and sometimes even the choice of the places where to work and live. In addition, companies belonging to states characterized by higher levels of taxation could be forced, in order to increase their attractiveness, to increase wage levels to compensate for the higher tax burden on income. This paper investigates whether differences in income taxation across US States affect the choices of the state (and consequently of the team) where to play for National Basketball Association players and whether different teams adjust their wages in relation to the state tax. The results obtained (probably due to the presence of salary cap) show that taxes do not affect neither the determination of wages or the choices of players.

It is likely that players choose where to play considering the probability of winning the championship and the socio-economic characteristics of the cities.

JEL Classification Codes: H24, H26, H31, H71, J31, Z22, Z28

Keywords: Taxation¸ Workers' Mobility, Sports Economics, Wage Determination

2.1 Introduction

The relationship between top income taxation and economic growth is a hot topic on the roundtables of policymakers and academics. Rigorous recent studies demonstrate the existence of a quadratic and non-monotonic relationship between them (Milasi and Waldman, 2017).

The vast majority of politicians consider the reform of the country's taxation system as a crucial point of their electoral campaign. For example, recently the concept of Flat Tax has become very famous in Italy and other European states, or, in the United States, Donald Trump focused in 2016 his electoral campaign for the presidency on tax reform[8], with the aim of reducing taxes for businesses and wealthier people, in order to produce an increase of economic growth[9].

In public economics, the theory of optimal income taxation studies taxation systems in order to minimize bias and economic inefficiencies. The introduction of a tax always leads to a change in the economic behaviour of individuals. For example, a tax on an asset causes a decrease in demand and a shift to other substitutes goods. Government revenue does not compensate enough for the loss of consumer surplus. This concept is called deadweight loss. In general, it is necessary to consider the problem of optimizing the tax burden in order to maximize GDP and at the same time to minimize the public deficit.

Each nation applies a different taxation system and a different withdrawal method to finance public spending. However, it is possible to compare the different "apparent tax burden" by comparing each country's total tax revenue to its gross domestic product (GDP)[10].

In western countries, in the last decade, the rates and brackets of personal income tax have been modified several times with the aim of reducing the number of brackets and lowering the marginal rate on higher incomes. This question is at the center of a heated debate whose fundamental issue is whether and how individual behaviours are modified by taxes. Those

[8] The President of the United States signed the tax cuts and jobs act in December 2017, two days after congressional approval.

[9] Trump argues that tax cuts for these categories of the population would have brought benefits even for the less affluent categories.

[10] Data published on Ocse’s website.

in favour of reducing taxes argue that the tax system has a negative effect on growth while on the contrary others argue that tax cuts have not affected the growth and behaviour of agents.

In recent times, the political discussion about the presence of tax progressivity has been linked to numerous studies on mobility response to taxation for several reasons: first, because mobility driven by top employee incomes increases the cost of tax efficiency and reduces governments' ability to redistribute and secondly, because there is no tax cooperation between countries (Kleven et al., 2019).

An important question is to determine whether people choose places in response to these tax differentials, thereby reducing the ability of local and national governments to redistribute income and provide public goods. Due to globalization and reduced mobility costs, it has become increasingly important to pay attention to responses on mobility in the design of tax policy.

Mirrless (1982) was the first who managed to determine the influence that taxation has on work commitments. He notes that individuals determine the quantity and type of work provided in a rational way, maximizing their utility function. By including the individual choice of the amount of work in the aforementioned function, Mirrlees makes explicit the trade-off between equity and efficiency. Equity, or the redistributive purpose, is obtained through a particular definition of the function of social well-being, intended as a function of individual utility levels, and by assigning different weights to the utilities of the subjects of the population. The “Welfarist” nature of this approach is therefore underlined. Efficiency, on the other hand, is pursued "by analyzing the effects of the tax system on job supply decisions"[11].

Considering the state migration, the results obtained in the literature are not univocal (Giraldo, 2017). Feldstein and Wrobly (1998) suggested that state and local governments cannot redistribute income. Considering that workers can avoid unfavorable taxes by migrating to location that offer more favorable tax conditions, this leads to an increase in wages in high tax states and a decrease in wages in low tax states (gross wages adjust rapidly to the changing tax environment). Thus, states cannot redistribute income for a

[11] Migration and Optimal income taxes (1982) published by Journal of Public Economics

period of even a few years. In contrast, these results were refuted by Young and Varner (2011).

Furthermore, for a long time it was not possible to produce relevant empirical evidence due to two important empirical challenges (Kleven et. al 2019): data challenges (it is very complicated to collect data that crosses migratory processes and precise measures of tax rates in different places) and identification challenges (the difficulty of identifying a fiscal variation that is orthogonal to all the other variables that can influence the choice of location, such as conditions of the labor market, crime level, public goods).

In recent years, the literature has produced various contributions in which the impact of tax levels on individuals' decisions to choose a specific location is treated, using different tools suitable to bypass the problems previously exposed.

Several recent studies have attempted to study the international migration of top incomes workers. Schimdheiny (2006), Schimdheiny and Slotwinski, (2018), Martinez (2017) presented empirical evidence regarding Switzerland (one of the countries with the best taxation levels in the world), emphasizing that top earners are strongly mobile within cantons.

The approach of our work follows what is described by Klaven et al. (2013) and Akcigit et al. (2016). Klaven et al., analyze the effects of top tax rates on international migration of football players in 14 European countries in the period of time between 1985 and 2008. The authors show evidence of strong mobility responses to tax rates, with an elasticity of the number of foreign (domestic) players to the net-of-tax rate around one (around 0.15). Furthermore, in this paper the authors found evidence of sorting effects and displacement effects [12].

The majority of the most relevant contributions regarding the relationship between taxation and workers' mobility in the literature used data collected in the United States.

The Constitution of the United States of America provides for a balance of power between the two levels of government, the federal and the state. Each state has its own written constitution, its own system of government and its own code of laws. The differences between the laws of the various states can be significant, even in important matters such as

[12] Sorting Effects are a particular incentive system that attract workers with specific characteristics, while Displacement Effects determine that low taxes on foreigners displace domestic players.

property, health, education and criminal law. The U.S. tax regime hinges on taxation at the federal, state and local levels. State taxes and taxes of some municipalities such as, for example, the city of New York must be added to federal taxes.

It is important to underline how tax rates vary a lot from state to state and these differences are very significant especially for those who are high income taxpayers (Moretti and Wilson, 2017).

In particular, there are seven states defined as "no income tax", because the state tax rate is equal to 0: Alaska, Florida, Nevada, South Dakota, Texas, Washington, and Wyoming. In addition, New Hampshire and Tennessee don't have a state income tax too, but they do tax interest and dividends at 5% and 2%[13]. These states are characterized by the presence of income taxes only at the federal level.

Akcigit et al. (2016) study the effect of top tax rates on “superstar” inventors’ international mobility since 1977, using panel data on inventors from the US and European Patent Offices. The inventors' choices were significantly influenced by the different measures of tax rates. In particular, the authors identified that the elasticity to the net of tax rate of the number of domestic superstars’ inventors was around 0.03, while for foreign inventors it was around 1.

Feldstein and Wrobel (1998), considering a sample of full time workers, have shown that wage levels change fully offset tax changes across states. Young et al. (2016) using data between 1999 and 2011 showed that millionaires are moderately mobile within the country.

The most significant impacts were identified by Moretti and Wilson (2017). In their paper, the economists quantify how sensitive is migration by star scientists to changes in personal and business tax differentials across states, showing that the long-run elasticity of mobility relative to taxes is 1.8 for personal income taxes, 1.9 for state corporate income tax, and 1.7 for the investment tax credit.

Previous works on the effect of the states' differences in top marginal income tax rates on sports salaries and migration in the United States show smaller effects. Alm et al. (2012) considered free agents in Major League Baseball and established that an increase in the top marginal tax rate of 1 percent augmented a free agent's salary by $21,000 to $24,000. Furthermore, in a paper that clearly measured the effects of migration on income tax rates,

[13] Bureau of the Fiscal Service (www.fiscal.treasury.gov)

Kopkin (2011) found a behavioral response in basketball players to an increase in tax rates: teams with higher state and municipality top marginal income taxes had a lower average skill of signed free agents.

The paper tries to verify if there is a relationship between the level of income taxation applied in a particular place and the behavior of taxpayers. The aim is to investigate migration responses to the net taxations in US States, exploiting sport as a vehicle for the analysis. In particular, the attention of the paper is addressed to the players of the National Basketball Association, in the period of time between the seasons 1995/1996 and 2011/2012. To achieve the goal, I use two econometric models. In the first specification I consider the model proposed by Kopkin (2011) in order to investigate whether the players' wages are affected by the level of state taxation. In the second specifications, I use a multinomial logit estimation to be able to derive players' elasticity of migration with respect to the net taxation rate (Kleven et al., 2013) (Giraldo, 2017).

Furthermore, before going into the analysis in more depth, I provided a basic evidence in order to emphasize the role that taxation has on the probability that a player decides to change state, using a Linear Probability Model and a Conditional Logit Fixed Effect Model.

The decision to use data from the NBA is motivated by two reasons. First, the reliability and quality due to the robust sources used; second, the league structure that allows players to move from one state to another year by year, respecting the determinants of the salary cap. In fact, this paper exploits this system, in relation to the effects that it can induce in the choices of the players by the teams, starting from the assumption that there are spending limits for the construction of each roster.

Moreover, I have considered two different time windows with regards to the analysis through the multinomial logit model, taking into consideration the definitive expansion that the league has had since the 2005/2006 season.

The paper is organized in the following way. Section 2 describes briefly the NBA Championship and discuss the relationship between NBA and taxation. In Section 3, I determine the model for player migration in National Basketball Association. In Section 4, I describe the dataset and present the descriptive statistics. In Section 5, I show the different empirical strategies used to study the goals of this paper and I illustrate the main results. Section 6 concludes.

2.2 The relationship between NBA and Taxation

The American National Basketball Association (NBA) was founded in 1946 by the owners of the largest American arenas, with the name of BAA (Basketball Association of American). In 1949 through the union of BAA and other minor leagues, the NBA was born with 17 participating teams. Over the years there were several relocations and changes that led to the creation of a championship with 30 participating teams (29 American and one Canadian), divided into two distinct conferences (eastern and western), according to a geographical criterion, each of which is made up of 3 divisions and each division is composed of 5 teams. Since 2004 during the season there are three different phases: Regular Season, Play-Off and Finals. The NBA Finals represent the last act of each season, decisive for decreeing the winning team of Larry O'Brien Trophy.

The roster of an NBA franchise generally consists of a minimum of 12 to a maximum of 15 players, officially under contract with the team. However, there are some exceptions and limitations.

The season starts between October and November, with each team playing 82 games during the regular season.

The salary cap is the budget that each team has available for their players' contracts, and is determined by the NBA based on the total revenue of the league. The cap for this season - and valid for all 30 teams - is set at $ 115 million, which represents the first separation threshold between the teams under the cap (and therefore has the greatest possible room for manoeuvre) and those above the cap (which have a number of restrictions on the buying and selling of players). The cap of each team, of course, is calculated by adding up the contracts of all the players. Unlike other American sports, such as hockey, the NBA cap is defined as soft, because teams can "push it" to sign their players through a series of rules. Basically, NBA teams can spend as much as they want for their players' contracts, but there is at least one further fundamental threshold that advises against the reckless accumulation of contracts: the luxury tax (set at 132 million in the last season): if the sum of the players' salaries exceeds that limit, the teams must pay an additional "luxury tax" to the league which is then redistributed to the teams that remain under the cap. There are two further thresholds: the salary floor, which provides that each team must spend at least 90% of the cap for its players (otherwise they must redistribute the deficit to the players under contract) and the apron, instead, placed just above the Luxury Tax (4 million for the season, therefore at 136) which represents an insurmountable limit for teams under the luxury tax but above the cap that use certain exceptions.

At this moment, there are three systems through which teams have the opportunity to recruit players: the draft, trades and the free agency.

Annually, like the other American sports characterized by the presence of the franchise-teams, also in the NBA in the month of June the draft lottery is carried out. It is an event that allows each team to choose, following certain criteria, new players that are at least 18 years old, coming from colleges or other international organizations to be included in their roster. The American sports system fixes that the training of young athletes takes place directly in high schools and colleges, excluding the existence of “youth teams”, determining that the only great possibility that teams have to ‘pick’ the young players with the best qualities is through this lottery. The first 14 choices belong to the teams that in the previous season did not have the strength to qualify from the regular season to the playoffs, with an order of choice determined through a lottery that follows certain rules. All players selected for the Draft are considered Rookies (players with the first year of experience in the NBA) and for selected in the first round there is a specially dedicated bargaining system: “the Rookie Scale Contract”.

The trades represent the exchanges of players that are carried out between the various clubs that create very proper "packages" composed, in some cases, also by cash and future choices in the draft.

The third way that determines player mobility is the free agency. The free agent is a player who is not under contract with any team and therefore can sign with anyone.

There are two different categories of free agents: unrestricted free agent and restricted free agent.

Unrestricted free agents are those players without a team (either because they have not been renewed by their previous team at the end of the contract or because they have not personally chosen to renew). These athletes are free to discuss with all the teams and subsequently decide with whom signing a contract. Restricted free agents are players who are free to request offers from other teams but who are obliged to have to re-sign with their last club if the latter matches any foreign offers.

Players who are not chosen in the annual draft, from college, are considered unrestricted free agents and are free to sign with any team.

Furthermore, teams cannot freely offer a figure to their players, in fact, there is an artificial limit to what a player can earn (maximum wages). This limit is directly linked to the salary cap, since a player can earn a predetermined percentage of the cap based on years of experience in the League: for players who have played for up to 6 years in the NBA, the maximum limit for the first year of the new contract is tied to 25% of the cap; for players with experience between 7 and 9 years, the limit is 30% of the cap; for those with over 10 years of experience, it is 35% of the cap. However, there are two ways to "advance among the brackets": if a player in his first four years of career is voted for an All-NBA quintet, he is twice nominated as All-Star Game holder or wins the MVP title, can access to the second salary bracket (therefore could reach 30% of the cap despite not having the experience).

In the United States there is an additional tax, the jock tax. It represents a tax that professionals are required to pay on all fees that are collected in a different state than that of residence.

It is very important to consider that there are very important differences, in terms of wages, between the different NBA players. For this reason, graphic evidence is shown about the distribution of players' salaries. In Figure 1 we show the long right-tail in the distribution of salaries, that is, the right skewness. Figure 2 shows that average salary earned by each decile: it is immediately clear how much higher are the salaries earned by the top deciles with respect to lower deciles.

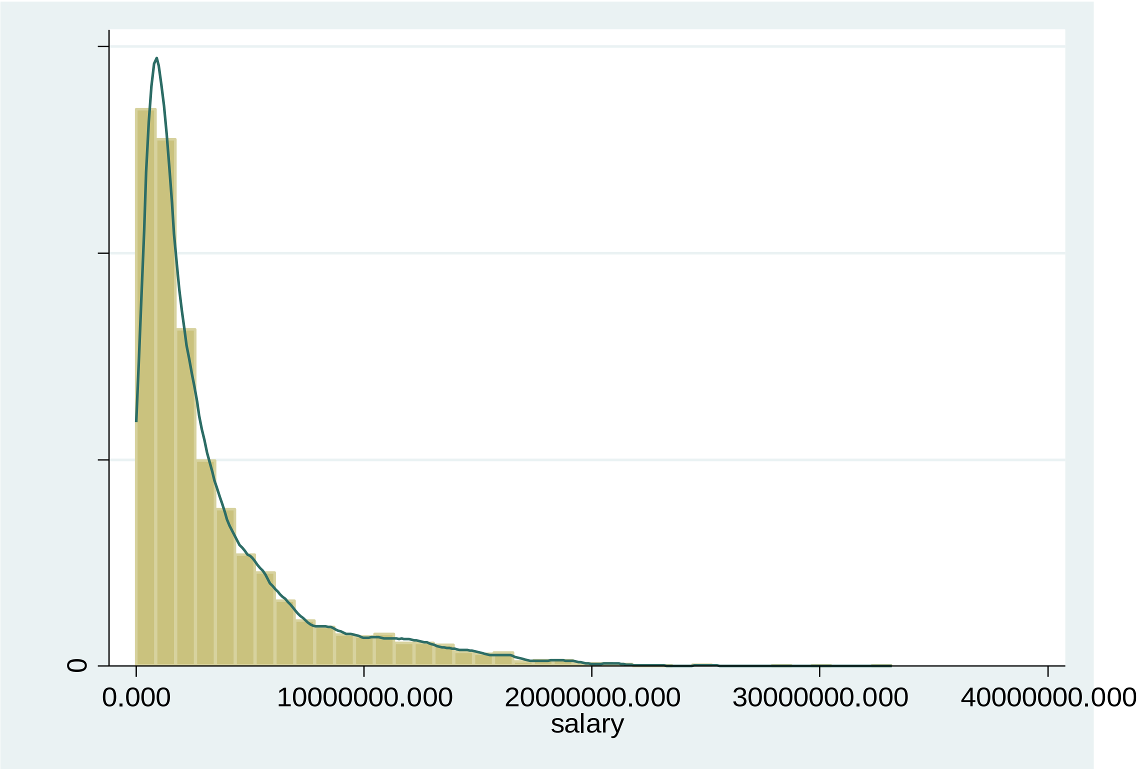

Figure 2.1: The skewness in the distribution of salaries

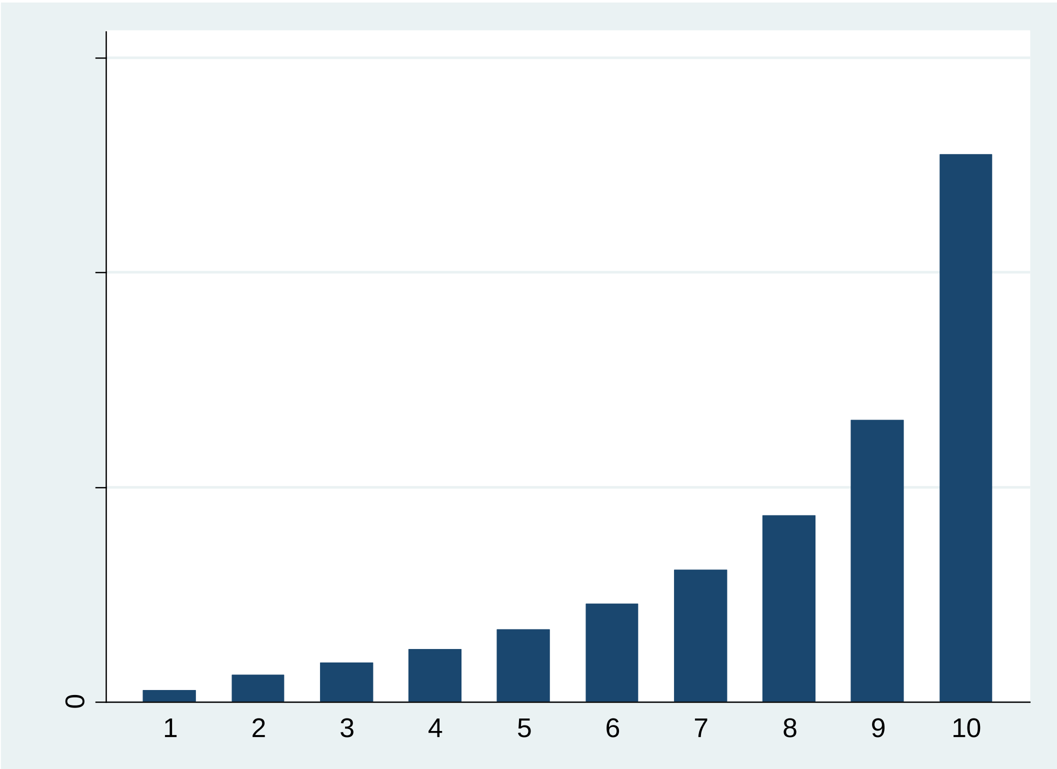

Figure 2.2: Mean of salary by decile

2.3 The Theoretical Framework

The theoretical framework used to determine the choice of players about the state in which they will decide to play is based on the empirical evidence obtained from Kleven et al. (2013) and Akcigit et al. (2016). Both studies used the same model to test the migration of footballers and scientists from one country to another. There are S teams, labeled as $s \in$ [1,…S]. Each player i earns a salary w_{ist}. At time t, players in each team must pay a tax rate τ_{st} on their total income. The economic utility received by player after taxation is equal to (1-τ_{st}) w_{ist}.

Considering an increasing function of a player's after tax income, the total utility from choosing a team *s* at time *t* for the player *i* is given by:

$$U_{ist} = u\big((1-\tau_{st})w_{ist}\big) + \varphi_{ist}$$

In previous works, Kleven and Akcigit consider the decisions of individuals between different countries, in our model the players choose between different American States.

In this model, φ_{ist} represent the different preferences that an individual has, regardless of the economic aspect. It is important to underline that the intrinsic characteristics of the team (habit of winning, record of victories in the previous season) or of the city (number of inhabitants, level of GDP per capita, crime rate) could affect the decision of players.

In this scenario, considering the absence of costs for moving, team *c* will be chosen in the period *t* if and only if:

$$u\big((1-\tau_{ct})w_{ict}\big) + \varphi_{ict} \geq u\big((1-\tau_{st})w_{ist}\big) + \varphi_{ist} \ \forall\ s \in S,\ s \neq c$$

Furthermore, considering the job demand, each team intends to maximize the sum of the abilities (e) of its players, subject to a budget constraint C, given by the level of the salary cap in NBA.

Hence, considering that a free agent *i* receives, at time *t*, two different offers: w_{1t} from team 1 (belonging to state A) e w_{2t} from team 2 (belonging to state B), with A≠B. In the case of φ_{i1t}=φ_{i2t}, free agent will accept the offer from team 1 if $w_{1t}(1\text{-}\tau_{1t}) \geq w_{2t}(1\text{-}\tau_{2t})$.

Additionally, assuming that A=B (two teams belonging to the same state, therefore $\tau_{1t} = \tau_{2t}$) e φ_{i1t}=φ_{i2t}, free agent will choose to play in team that will offer the largest w_t. Consequently, in the case of w_{1t}=w_{2t} and φ_{i1t}=φ_{i2t}, free agent will sign with team 1 if (1-τ_{2t}) ≥(1-τ_{1t}).

Rationally, the teams have the task of compensating any higher tax taxes charged in the state of belonging with higher wages. In the NBA, as in other American professional sports, the presence of salary cap (especially the luxury cap) prevents the tax differences from being paid perfectly through higher wages, for this reason it could happen that players with more skills would tend to end up in teams located in states with lower tax rates (Giraldo, 2017).

It is also very important to clarify a crucial aspect: in light of the fact that the NBA seasons do not coincide with the calendar year (the championship generally starts in October, the free agency period mainly occurs in August), basketball players choose their team considering the tax levels in force in the same calendar year as their choice.

2.4 Data and Descriptive Statistics

Empirical analyses required the collection of data regarding salaries, physical characteristics, biographical characteristics and the main technical statistics of the players of the American professional league. The present work exploits a wealth of data. In fact, we proceeded to acquire data relating to the NBA superstars starting from the 1995/1996 season up to the 2011/2012 season, as many as 17 seasons that allowed to carry out different analyses drawing on a dataset consisting of 1128 players who succeeded each other in the American arenas over the observed years, leading to the construction of a balanced panel.

Over the years covered by the analysis, the league has had a phase of expansion. In 1995 the number of teams went from the initial 27 to 29: with the extension of the NBA in Canada, were born the Toronto Raptors. In 2004 the NBA reached 30 franchises, with the birth of the Charlotte Bobcats, and continues to evolve as one of the most important and best organized professional sports leagues in the world. Undoubtedly, a decisive contribution for the establishment of the dataset was provided by the rigidity and clarity that distinguish the league regulations. In fact, all the franchises are obliged, due to salary cap, to publish and systematically update the salary data of each basketball player regularly under contract, allowing everyone, unlike other 'European' sports, to have a large amount of information and data available in real time. The source used to obtain information on salaries and main performance characteristics was the website www.basketball-reference.com. Biographical information was acquired from the official website of the league www.nba.com. To verify the robustness of the data collected, we carried out a further check considering the basketball datasets provided by the website www.espn.com.

To comply with the empirical evidence, it was necessary to identify, among those available, a performance variable capable of uniquely representing the individual performance of the players and being suitable for carrying out comparisons between them considering the different seasons of the dataset. For this reason, I have chosen PER. PER is a player's efficiency rating (the sum of all the positive and negative contributions of a player, which takes into account the minutes played and the performances of all the players in the League) created by the statistician John Hollinger of ESPN.

In this sample the average PER value is 13.35. The highest season values were achieved by LeBron James in 2009 (31.7), who won the MVP award.

2.4.1 Tax Rate Data

In order to identify information about the different levels of taxation, we used the same data shaped specifically by Moretti and Wilson (2013), for the paper "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists".

Using the NBER's TaxSim simulator and the World Top Income Database, the authors obtained information on income tax rates by states and year for a hypothetical scientist with a salary and capital gains income belonging to the various national income percentiles. First, the authors collected information, for each year from 1977 to 2010, on the 95th, 99th, and 99.9th percentiles of the national income distribution (separately for salary and capital gains) from the World Top Incomes Database (Alvaredo, Atkinson, Piketty & Saez 2013) and data on the 50th percentile from the Congressional Budget Office (CBO). Subsequently, they entered these data into the simulator, considering the following assumptions: the taxpayer was a married joint filer, less than 65 years old, had zero dependent exemptions, zero childcare expenses, no other sources of income, and zero itemized deductions other than the deduction for state income tax payments.

For more information about the construction of the dataset carried out by Moretti and Wilson, it is possible to consult the online appendix on the page https://www.aeaweb.org/content/file?id=4801.[14]

In literature, the studies carried out regarding the relationship between taxation and worker migration have not clearly established which variable between average tax rate and marginal tax rate is more correct to use to obtain more efficient estimates.

The average tax rate (ATR) is the total amount of tax divided by total income. The marginal tax rate (MTR) is the incremental tax paid on incremental income (if an individual were to earn an additional $ 1,000 in wages on which he paid $ 100 of income tax, the individual's marginal tax rate would be 10%). Contrary to the MTR, the ATR is not a statutory rate, but obviously they are strongly correlated (Moretti & Wilson, 2013). In this work, I tried to take advantage of the state level changes in the average tax rate considering a hypothetical worker at 999th, 99th and 95th percentiles of the national income distribution, obtaining very similar results for each level. The assumptions, highly realistic, are that the salary of NBA players is in the top 5 percent and that the latter choose the team in relation to the average tax rates, not the marginal tax rates. Federal taxes and state rates are included in my ATR measure.

2.4.2 Additional Data

In order to control for other factors, I collected data about the particular metropolitan area of each team. Specifically, I collected information about the level of GDP per capita and the number of inhabitants, directly from the Bureau of Economic Analysis. In addition, in

[14] For Canada, I took data from the website of the Federal Revenue Agency.

order to obtain more robust and efficient estimates, I proceeded to build a variable capable of capturing the number of games won in the previous season for each team and the number of championships won in their history. This sport information was obtained directly from the league's website www.nba.com. Probably, the players could decide to choose a certain team also in relation to the blazon or the chances of victory (skilfully captured by the performances recorded in the previous seasons).

2.4.3 Descriptive Statistics

The dataset consists of information on over 1000 players in the 17 seasons considered.

For each athlete I obtained data on salary, biographical characteristics and performance measures (both standard and advanced). For each state with a team participating in the championship, I collected information on the different levels of average tax rates (considering the different percentiles), on the number of inhabitants of the metropolitan area of the city of reference and the corresponding level of GDP per capita.

Descriptive statistics are reported in Table 2.1.

The average salary is equal to $ 395379, while the median salary is lower (1,995,676). The age is 26.75. On average players are 200 cm tall and weigh around 98 kg. In the period of time considered, athletes play about 57 games over an 82-game season, recording an average PER of 13.35 (practically equal to the median PER equal to 13.3). In addition, on average, they scored 504.377 points, providing 13.201 assists. The average levels of Real Gross Domestic Product per capita achieved in the states considered was around $ 45736, while on average the states considered had a population level of about 14 million. The average tax rate for a hypothetical payer is 34.9% at the 999th percentile, 30.3% at the 99th percentile and 11.8% at the 50th percentile.

On average, the players in the sample have won about 2 NBA titles.

In this paper, I only consider the states that within them have at least one team registered in the NBA championship, in the analysed period. To provide further information, Table

2.2 lists the NBA teams by state of origin. The most represented states are California (4 teams: 2 in the city of Los Angeles, Sacramento, Auckland) and Texas (3 teams: Dallas, Houston and San Antonio).

Table 2.1: Descriptive Statistics

	mean	SD	min	max
Year	2003.224	4.871	1995	2011
Age	26.753	4.175	18	43
Games Played	57.689	23.256	1	82
Games Started	30.134	30.482	0	82
Minutes Played	1430.764	919.713	1	3464
PER	13.351	5.482	-90.6	90.3
Assist	13.201	9.408	0	76.1
Steals	1.667	.835	0	12.1
Blocks	1.599	1.949	0	77.8
Turnovers	14.322	5.277	0	100
Field Goals	193.395	186.601	0	1251
Field Goals Attempted	428.861	397.113	0	2457
Points	504.377	487.311	0	3033
Salary	3953279	4043647	2706	6298258
GDP	45736.79	9511.111	22335	65746
Population	27019.721	20359.931	2014177	71545.178
ATR_p50	.118	.019	.0782	.168
ATR_p95	.222	.028	.164	.278
ATR_p99	.303	.028	.243	.345
ATR_p999	.349	.0342	.277	.415
Height	200.819	9.511	160	231
Weight	98.735	12.695	60	147
Champs	1.978	3.833	0	17
Observation	5293			

Table 2.2: NBA teams and States

State Name	NBA Team (S)
Arizona	Phoenix Suns
California	Golden State Warriors Los Angeles Clippers Los Angeles Lakers Sacramento Kings
Colorado	Denver Nuggets
District of Columbia	Washington Wizards
Florida	Miami Heat Orlando Magic
Georgia	Atlanta Hawks
Illinois	Chicago Bulls
Indiana	Indiana Pacers
Louisiana	New Orleans Pelicans
Massachusetts	Boston Celtics
Michigan	Detroit Pistons
Minnesota	Minnesota Timberwolves
New York	Brooklyn Nets New York Knicks
North Carolina	Charlotte Hornets
Ohio	Cleveland Cavaliers
Oklahoma	Oklahoma City Thunder
Oregon	Portland Trail Blazers
Pennsylvania	Philadelphia 76ers
Tennessee	Memphis Grizzlies
Texas	Dallas Mavericks Houston Rockets San Antonio Spurs
Utah	Utah Jazz
Wisconsin	Milwaukee Bucks

Table 2.3 describes the average levels of ATR obtained in the states considered in this article. Considering the different percentile levels, the highest values are recorded in California and North Carolina while in Florida, Tennessee, Texas and Washington the lowest values, reflecting the differences due to the presence or absence of state taxes.

Table 2.3: Mean of ATRs by US State

STATE	ATR_p95	ATR_p99	ATR_p999
Arizona	0.221	0.302	0.355
California	0.243	0.331	0.381
Colorado	0.228	0.306	0.355
District of Columbia	0.241	0.330	0.379
Florida	0.185	0.259	0.311
Georgia	0.239	0.314	0.359
Illinois	0.215	0.290	0.338
Indiana	0.216	0.291	0.339
Louisiana	0.209	0.291	0.332
Massachusetts	0.24	0.312	0.359
Michigan	0.226	0.302	0.347
Minnesota	0.245	0.326	0.367
New York	0.244	0.324	0.368
North Carolina	0.256	0.324	0.381
Ohio	0.231	0.314	0.363
Oklahoma	0.241	0.314	0.357
Oregon	0.259	0.336	0.336
Pennsylvania	0.214	0.288	0.336
Tennessee	0.183	0.260	0.311
Texas	0.184	0.260	0.311
Utah	0.238	0.313	0.358
Wisconsin	0.248	0.319	0.36
Total	0.224	0.302	0.349

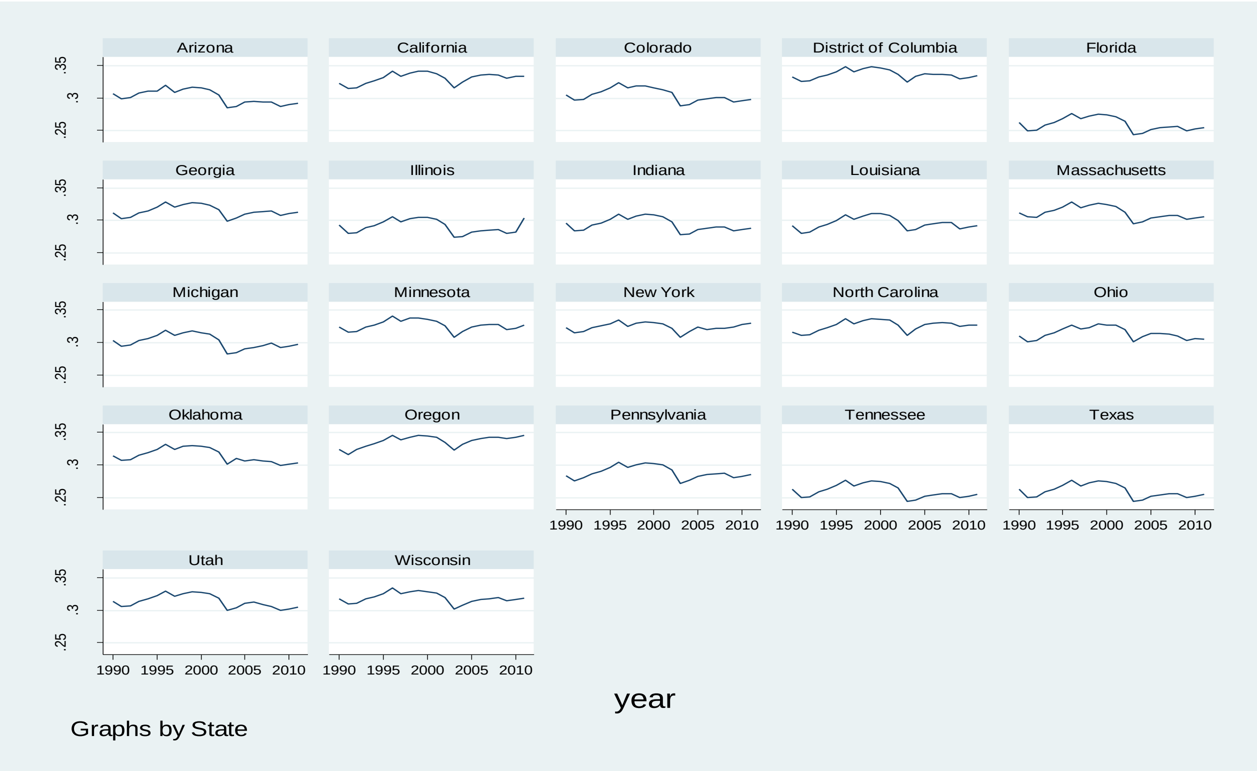

Figure 2.3: The relationship between ATR_p99 and Year by State

Figure 2.3 allows us to graphically check the trend of the different levels of taxation in the period of time analysed, considering the different states. In particular, we can see how the lowest levels of taxation were logically recorded in the states characterized by the presence of taxes only at the federal level (Florida, Tennessee and Texas).

2.4 Empirical Strategy

The goal of this work is to try to estimate the effect of the variation of the different levels of tax rates on the choice of the team by a NBA player. To achieve this, I considered two different econometric models, both already presented by Giraldo for a similar work on MLB (2017).

2.4.1 A preliminary analysis

In order to be more clear, the effects of the relationship between players' mobility and the different levels of taxation across States, I decided to use additional models (linear and non-linear) to verify whether the probability of an individual decision to move from a state to another may be conditioned by the different level of taxation.

Considering the dummy variable Change equal to one if the player has decided to move from one state to another as our dependent variable, such that:

$$\text{Prob (Change=1|x) = F (x, } \beta\text{)} \qquad (1)$$

where x is the vector of the explanatory variables and β is the vector of coefficients that determine the change in x on the probability.

Basically, the simple model I estimate is the following:

$$Prob\ (Change_{it} = 1|\text{x}) = \beta_0 + \beta_1 NoStateTax + \beta_i X_{it} + \varepsilon_{it} \qquad (2)$$

β_1 is the coefficient of the main variable, referring to taxation levels, X_{it} is the vector of the other explanatory variables (individual characteristics, sports performance, characteristics of the states) and ε_{it} is the error term , I use a dummy variable defined NoStateTax which is equal to 1 for the three states in our sample without state income tax (Florida, Texas, Tennessee) and 0 for all other states.

In order to correctly estimate this model, I used two different strategies: an estimate with the OLS to provide a first basic evidence (LPM, linear probability model) and an estimate through the logistic distribution, determining the following logit model:

$$\text{Prob (Change=1|x)} = \frac{e^{x_i\beta}}{1+e^{x_i\beta}} = \Lambda\ (x_i\beta) \quad (3)$$

where Λ (.) represents the logistic cumulative distribution function.

Moreover, exploiting my panel data, in order to take into account, the unobserved heterogeneity, I use a conditional fixed effects method.

Following Allison (2009), there are two conditions that I had to consider:

a) The dependent variable must have at least two measurements for each player;

b) The independent variables must vary a lot over time between players.

Clearly, with this estimate we tolerate the loss of observations referable to all those players who have never moved from their home state, during the analyzed period. However, the choice fell on these tools, as other methods, eg. random effects, will suffer from omitted variable bias; fixed effects methods help control omitted variable bias by having individuals act as controls (Allison, 2009).

The results obtained are shown in Table 2.4. Column 1 shows the results of the simple linear probability model while column 2 shows the marginal effects of the logit model. The evidences obtained are quite similar, in fact we can verify the existence of a quadratic relationship between the dependent variable and the age variable (inverted U-Shape): at the beginning of their career players prefer to move more frequently, on the contrary in the last years they decide to stay in a team to compete for the championship. The performance variables of the previous season show their importance: the PER, the games in starting lineup and the victories obtained have an inverse effect on the athlete's probability of moving, therefore lower performances, fewer chances to play and poor results invite players to change teams. The third column shows the last specification, having the conditional fixed effects available, bringing small and significant variations with respect to what is expressed by the previous columns. The quadratic relationship with the age variable resists, while the only significant variable regarding the previous season's performance is the number of victories that the previous team recorded in the league. In fact, the lower the number of victories that a basketball player has achieved in the previous season, the higher the probability of the player to change teams. This leads to the assertion of another important aspect that can push an athlete to move: the ambition to win.

Our variable of interest in this estimate assumes statistical significance at 10%. In particular, the possibility of achieving lower levels of taxation in other states increases the likelihood that a player decides to change state.

Table 2.4: The preliminary analyses

	LPM (1) Change	Logit (2) Change	Conditional FE (3) Change
Age	0.146***	0.161***	0.698***
	(0.018)	(0.019)	(0.137)
Age2	-0.002***	-0.002***	-0.010***
	(0.000)	(0.000)	(0.002)
Previous Season GS	-0.001***	-0.001***	0.001
	(0.000)	(0.000)	(0.002)
Previous Season PER	-0.007***	-0.007***	-0.010
	(0.001)	(0.001)	(0.009)
Δ Wage	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)
GDP	0.000***	0.000***	0.000
	(0.000)	(0.000)	(0.000)
Population	-0.000**	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)
Previous Season Win	-0.323***	-0.321***	-1.465***
	(0.044)	(0.043)	(0.268)
Champs	-0.002	-0.002	0.006
	(0.002)	(0.002)	(0.012)
No_State_Tax	0.038	0.037	0.216*
	(0.029)	(0.029)	(0.113)
ATR_p99	0.273	0.268	
	(0.414)	(0.410)	
_cons	-1.744***	-11.535***	

	(0.247)	(1.243)	
N	4891	4891	3883
adj. R^2	0.085		

This table considers the entire sample, except for the third column in which I don't have the observations referable to all those players who have never moved from their state. This table shows two different models. In the first column I show a LPM model with standard error robust to heteroschedasticity. The second and the third columns consider the logistic distribution. In particular, in the third column I use the conditional fixed effect, considering player as panel variable. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at 1,5 and 10% level

2.4.2 The relationship between taxes and wages

The first model was proposed by Kopkin (2011) for the NBA. For all free agents I have estimated the following equation:

$$W_{ist} = \beta_0 + \beta_1 \tau_{st} + \beta_2 V_{it} + \lambda_t + \varepsilon_{ist}$$

W_{ist} represents the salary that player *i* receives at time *t* from team *s*, τ_{st} represents the tax rate paid by the hypothetical payer belonging to the 99th percentile in state *s* at time *t*. The goal of this model is to try to test whether the players' wages are affected by the level of state taxation. In other words, if the teams compensate for the higher value of the tax rate in the state of affiliation by giving a higher salary to the players that they decide to hire.

Clearly, I included controls in the model. V_{it} is the vector representing the characteristics of players (both measures of performance and biographical); λ_t are time fixed effect.

Considering the seasons present in the dataset, the time fixed effects are useful in order to consider the undetected wage changes due to the economic trends and renegotiations of the Collective Bargaining Agreement. The reference coefficient of the model is β_1. In the event that β_1 is positive, the teams belonging to the states with higher tax rates offer (or are forced to offer) a higher level of wages to compensate for the greater tax expenditure by the players. The results of this model are shown through the use of Table 2.5. For the purposes of the estimate, I have considered only free agent players, having more than 5 years of experience in the league to eliminate the problem of the rookie scale contract[15].

In the first column, I showed a first estimate concerning only the value of ATR_p999 and age. Subsequently, following the determinants of the literature, I inserted the variable age squared in the equation to try to capture a possible non-linear relationship with wages. The results obtained show the 'famous' U-Shape relationship between the variables, highly significant. In addition, I use as proxy of the performances of the players a single variable: the PER index. From column 2, we can underline how much the performance affects the salary of the players. In particular, a unit increase in the PER index leads to an average wage increase of approximately 4.7%.

In the third column I added the time fixed effect, useful in order to consider the undetected wage changes due to the economic trends and renegotiations of the Collective Bargaining Agreement.

The complete model is represented by the fourth column. In this estimate, I tried to check for the different roles that players can play during the game (Role Controls)[16].

Table 2.5: The relationship between taxes and wages

	ln_wage	ln_wage	ln_wage	ln_wage
ATR_p999	0.064	1.015	1.188	1.073
	(0.819)	(0.858)	(0.847)	(0.836)
Age	0.016	0.617^{***}	0.939^{***}	0.952^{***}
	(0.010)	(0.133)	(0.122)	(0.123)
Age2		-0.010^{***}	-0.012^{***}	-0.013^{***}
		(0.002)	(0.002)	(0.002)
PER		0.047^{***}	0.044^{***}	0.045^{***}
		(0.009)	(0.009)	(0.009)
Role Controls				YES
Time FE			YES	YES
FE		YES	YES	YES
N	1746	1746	1746	1746
adj. R^2		0.068	0.159	0.160

Robust standard errors clustered at team level in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at 1,5 and 10% level

[15] The first-round draft choices receive an assigned amount of salary, according to their draft position. These contracts are for two years, with a team option for the third and fourth seasons.

[16] Considering the five basic positions in basketball we use as roles: Point Guard, Shooting Guard, Small Forward, Power Forward and Centre.

In each estimate, the coefficient of ATR_p999 assumes a positive value, as desired, but is not statistically significant at 10%.

The conclusion of this model is that there is no statistical relevance to the decisions of teams in states with higher tax levels to grant higher salary levels in order to attract players by compensating for their higher tax losses.

For an additional check, I got similar results considering ATR_p95 and ATR_p99.

2.4.3 The Multinomial Logit analysis

The second model aims to determine the elasticity of the migration to the net of tax rate.

Following the approach of Kleven et al. (2013), the model is a multinomial discrete choice model. In particular, in this work I selected to use a McFadden's choice model (1974). Discrete choice models are used to describe, understand and predict the choices of individuals when the set of choices is made up of a finite number of possibilities.

They imply that for each individual i, a given level of utility is associated with each alternative j. The alternative j^* will be chosen if and only if the relative utility, among the set of choices, is superior. This utility can depend both on the characteristics (attributes) of the alternatives and on the socio-economic characteristics of the individuals. McFadden's Choice Model is a discrete choice model that uses conditional logit, in which the variables that predict choice can vary either at the individual level (in our case players), or at the alternative level (in our case teams).

From the theoretical model explained above, we have determined the following utility function:

$$U_{ist} = u\big((1-\tau_{st})w_{ist}\big) + \varphi_{ist}$$

Considering a logarithmic relationship, exploiting its properties, inserting the V_{it} characteristics of a player and the team fixed effects λ_s, we obtain the following fundamental equation in our work:

$$U_{ist} = \lambda_1 \log(1 - \tau_{st}) + \lambda_1 \log(w_{ist}) + \lambda_2 V_{it} + \lambda_s + \varepsilon_{ist}$$

The McFadden's choice model allows me to consider the counterfactual top average tax rate that player could have if they chose to play in the other 29 teams.

The problem with this kind of analysis is that we don't have the salary that the other 29 teams could offer to a particular player. In other words, for each player we only know the salary offered by his team. This problem in the NBA (in general, in American sports) is less important than in other scenarios, thanks to the presence of the salary cap and the limitations (as if they were real standardizations of the offers) imposed on the teams as regards the remuneration to be offer to different players in the league.

However, to further reduce the incidence of this problem, I thought it appropriate to present different types of specifications.

The last empirical step of the work concerns with the determination of the elasticity measure.

Considering the utility function explained above and $P_{ist} = \Pr(U_{ist} > U_{ict}, \forall c)$, that is, the probability that player *i* decides to play at time *t* for team *s*. Assuming that the error term ε_{ist} is type I extreme value distributed, the multinomial logit model can be estimated by maximum likelihood (Kleven et al., 2013). Hence, the estimate of λ_1 is useful for estimating the elasticity of the probability that an individual *i* will choose the team *j* at time *t* with respect to the net of tax rate (Kleven et al., 2013) (Akcigit et al., 2016) (Moretti et al., 2017) (Giraldo, 2017). In particular, a positive value of λ_1 implies that an decrease in the net of tax rate in a given state has a positive effect on a player's probability of choosing a specific team belonging to that particular state. Starting from λ_1 and taking advantage from the mathematical derivations of Kleven et al. (2013) and Moretti et al. (2017) it is possible to derive the standard formula of elasticity for the individual *i* of the probability of choosing team *s* at time *t* (P_{ist}).

$$P_{ist} = \frac{e^{\lambda_1 \log(1-\tau_{ist}) + \lambda_2 V_{ist}}}{\Sigma_g e^{\lambda_1 \log\left(1-\tau_{gst}\right) + \lambda_2 V_{igt}}} \quad \text{and} \qquad E_{ist} = \frac{\mathrm{d} \log P_{ist}}{\mathrm{d} \log(1-\tau_{st})} = \lambda_1 \left(1 - P_{ist}\right)$$

In addition, it is possible to determine team s level elasticities(E_s), such as the sum, weighted by P_{ist}, of the elasticities for all players for that team (Giraldo, 2017):

$$E_s = \frac{\lambda_1 \Sigma_i (1 - P_{ist}) P_{ist}}{\Sigma_i P_{ist}}$$

Finally, the average elasticity of the probability of choosing a team to the net of tax rate as the average weighted elasticities across the teams will be determined by the following equation (Kleven et al., 2013) (Giraldo, 2017):

$$\mathrm{E} = \Sigma_{s=1}^{30} \left(\frac{\lambda_1 \Sigma_i (1 - P_{ist}) P_{ist}}{\Sigma_{s=1}^{30} \Sigma_i P_{ist}} \right)$$

In order to estimate correctly through this model, we consider two different time windows, due to the addition of a team starting from the 2005/2006 season (Charlotte Bobcats).

The next table (2.6) shows the results in the period between the 1995/1996 season and the 2004/2005 season.

Table 2.6: Multinomial Discrete Choice Estimations 1995/1996-2004/2005

	(1)	(2)	(3)	(4)	(5)
Ln (1-ATR)	-0.853	-0.428	7.038	1.085	1.35
	(0.528)	(0.558)	(10.28)	(11.67)	(11.71)
Ln (GDP)		0.218**	0.354	0.707	0.680
		(0.0549)	(0.819)	(0.458)	(0.860)
Ln (Pop.)		0.0309	0.264	1.449	1.605
		(0.0232)	(1.047)	(1.283)	(1.286)
Prev.Season Win				0.137	0.157
				(0.108)	(0.174)
Team FE			YES	YES	YES
Players Characteristics					YES
Elasticity	-0.822	-0.412	6.773	1.045	1.388
Observations	73976	73976	73976	66164	66164
Cases	2642	2642	2642	2363	2363

The table shows a multinomial logit regressions considering the period of analysis 1995/1996-2004/2005. Robust standard errors clustered at individual level in parentheses. The symbols **, * indicate that coefficients are statistically significant, respectively, at 5 and 10% level

In the first column, I only used the net of tax rate as an independent variable. In the second column, I added the characteristic data of the metropolitan areas (GDP level by capita and population, both expressed in logarithmic terms). Next, to check the unobserved characteristics in teams that influence player's decision, I added team fixed effects.

In the last and more complete estimate, I also included the characteristics of the players (age and PER) and the percentage of games won by the teams in the previous season. In fact, players may decide to switch teams also in relation to their ambitions. Players who have not won championships in their career, may decide to choose the team that they think could maximize the chance of winning the championship.

Considering the first simple specification we find an unexpected and momentarily negative sign of the coefficient of ln (1-ATR).

By entering the data on the metropolitan area, the sign of the variable continues to be negative and no statistically significant, unlike the values of ln_GDP.

In column (3), I add the team fixed effects. The coefficient of the variable of interest reverses the sign, becoming positive, showing a high magnitude value (7,038), much higher than the estimates presented by Giraldo (2017) and Kleven et al. (2013). By inserting the characteristics of the players and the percentage of games won by the teams in the previous season, the coefficient of ln(1-ATR) continuing to show a positive sign and to be no statistically significant. We get similar results considering only free agent players in the sample.

Since 2005 the league has reached the current formation of 30 teams at the start of the championship, for this reason I decided to present the results only of this time window not subject to any expansion.

Table 2.7: Multinomial Discrete Choice Estimations 2005/2006-2011/2012

	(1)	(2)	(3)	(4)
Ln (1-ATR)	0.095	0.532	7.613	9.475
	(0.496)	(0.527)	(7.346)	(7.407)
Ln (GDP)		0.252**	1.936**	2.177**
		(0.091)	(0.938)	(0.950)
Ln (Pop.)		0.043*	4.153**	3.629**
		(0.025)	(1.728)	(1.762)
Previous SeasonWin				0.334*
				(0.189)
Team FE			YES	YES
Elasticity	0.092	0.512	7.322	9.113
Observations	59696	59696	59696	59696
Cases	2132	2132	2132	2132

The table shows a multinomial logit regressions considering the period between 2005/2006 and 2011/2012. Robust standard errors clustered at individual level in parentheses. The symbols **,* indicate that coefficients are statistically significant, respectively, at 5 and 10% level

Considering this more modern and representative sample of the current composition of the alloy, we can confirm the same empirical evidence obtained in the previous analysis regarding the variable of interest ln (1-ATR). In particular, no statistically significant relationship is highlighted between the net rate of taxation and the dependent variable. As expected, however, we observe the presence of positive sign coefficients in the columns representing the most complete models.

The important finding in this analysis is the relationship between the dependent variable and the other control variables. In particular, we can assert that the choice of the location by NBA players is positively influenced by the factors inherent in the metropolitan areas and the percentage of games won by a particular team in the previous season. This would lead to assert that players tend to choose a team taking into consideration the assets present in the teams and that a player's ambition about the chances of winning the championship plays a key role in the choice of location.

As for the coefficients of elasticity, being clearly linked to the values obtained by our variable of interest ln (1-ATR), we can assert the same conclusions expressed above. In fact, we observe the presence of positive coefficients, characterized by a high magnitude in columns (3) and (4), which are not statistically significant at conventional reference levels.

2.5 Conclusion and Further Research

In this work, I studied the effect of taxation on the national migration of basketball players belonging to the NBA, with the aim of inserting the paper into the flourishing literature that deals with analysing the connections between the tax systems and the labour supply.

Considering the chosen scenario of American basketball, a further objective of this work is represented by the analysis of the salary cap. Specifically, I checked whether the remuneration system envisaged within the NBA could be suitable for containing opportunistic behaviors, from a fiscal point of view, by the league's players in the choice of the team.

To achieve these objectives, I estimated two different statistical models, proposing different specifications to obtain robust and efficient estimates.

In the first model, I verified that teams do not adapt the wages of NBA players to the different taxes, considering the state in which they are located, demonstrating a first consequence of the presence of the Salary Cap (it does not allow an effective compensation of the higher tax rate paid by the players through a related increase in their wages). A key role in the determination of wages is given by the levels of performance that are achieved by the players, highlighting a sort of "wage meritocracy".

In the second model, I have implemented a multinomial discrete choice model (McFadden's Choice Model), in order to determine and analyse the choices of individuals when the set of choices is made up of a finite number of possibilities. In addition, this model allowed me to calculate different values of elasticity of NBA players to the tax rate.

Moreover, through this model (using different specifications), I did not obtain significant results in the relationship between the choice of the team and the levels of taxation present in the team's state, considering the two different time windows analysed. Overall, this paper allows to underline how players tend to follow other important reasons, in the choice of the team: personal ambitions and socio-economics characteristics of the cities.

In particular, the victories obtained in previous seasons represent an important element taken into consideration by the players in the decision to change team, signalling a strong intention to favor the sport aspect (the championship victory) over the fiscal aspect.

In light of the fact that in the 2011/2012 season a new and very rich collective agreement for players came into force which completely changed the economic schemes proposed in the league, it would be appropriate to implement this study, analysing the time span between 2011/2012 and 2019/2020 (sadly remembered as the season influenced by Covid-19), to study the targets of this paper in a very prosperous period for the league. Furthermore, a step forward in this research would be represented by the implementation of the analysis considering all the American sports that are characterized by the presence of the Salary Cap, to compare with greater force the conclusions drawn in this work.

2.5 References

Akcigit, U., Baslandze, S., Stantcheva, S., (2016), "Taxation and the International Mobility of Inventors", American Economic Review, 106 (10), 2930-2981.

Allison, P. (2009), "Fixed Effects Regression Models", Sage Publications.

Alm, J., Kaempfer, W. H., Sennoga, E. B. (2012), "Baseball Salaries and Income Taxes: The "Home Field Advantage" of Income Taxes on Free Agent Salaries", Journal of Sports Economics, 13(6), 619–634.

Feldstein, M. and Wrobel, M.V., (1998), "Can State Taxes Redistribute Income?", Journal of Public Economics, 68 (3), 369-396.

Giraldo D., (2017), "Taxation and Migration: evidence from Major League Baseball", University of Yale.

Kleven, H., Landais, C., Munoz, M., Stantcheva, S. (forthcoming), "Taxation and Migration: Evidence and Policy Implications", Journal of Economic Perspectives.

Kleven, H., Landais, C., Saez, E., (2013), "Taxation and International Migration of Superstars: Evidence from the European Football Market", The American Economic Review, 103 (5), 1892-1924.

Kopkin, N., (2011), "Tax Avoidance: How Income Tax Rates Affect the Labor Migration Decisions of NBA Free Agents", Journal of Sports Economics.

Martinez, I., (2017), "Beggar-Thy-Neighbour Tax Cuts: Mobility After a Local Income and Wealth Tax Reform in Switzerland", LISER, Working Paper Series 2017-08.

McFadden, D.L., (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in Paul Zarembka ed., Frontiers in Economics, Academic Press, chapter 4, pp. 105-142.

Milasi, S. and Waldmann R.J. (2017), "Top Marginal Taxation and Economic Growth", Applied Economics 50:19, pp. 2156-2170.

Mirrlees, J. A., (1982), "Migration and Optimal Income Taxes". Journal of Public Economics, 18 (3), 319-341.

Moretti, E., Wilson, D.J., (2017), “The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists”, American Economic Review, 107:17 pp. 1858-1903.

Schmidheiny, K., (2006), “Income Segregation and Local Progressive Taxation: Empirical Evidence from Switzerland.” Journal of Public Economics 90 (3): 429–58.

Schmidheiny, K., Slotwinski, M. (2018), “Tax-Induced Mobility: Evidence from a Foreigners’ Tax Scheme in Switzerland”, Journal of Public Economics, 167, 293–324.

Young, C., Varner, C., (2011), “Millionaire Migration and State Taxation of Top Incomes: Evidence from a Natural Experiment”, National Tax Journal, 64 (2), 255-284.

CHAPTER THREE

<u>THE IMPACT OF INCREASED PENALTIES FOR CAUSING ROAD ACCIDENTS AND DRIVING OFFENCES: EVIDENCE FROM ITALY</u>

Road safety is one of the most discussed issues among institutions and lawmakers. The main tools used to increase road safety are the various prevention campaigns and the introduction of stricter rules capable of guiding the behavior of individuals. Using data on a daily basis, we estimate the effect of the introduction of the "Vehicular Homicide Law" (VHL) on road accidents and on driving offences in Italy, through the use of a Regression Discontinuity Design, controlling for a series of variables such as holidays, unemployment rate, taxes on beer, police patrols, gasoline price. We observe the presence of an inverse and statistically significant relationship between the number of people sanctioned for a blood alcohol level higher than the amount that is legally allowed and the introduction of the new law (-2.628 persons sanctioned on a daily basis), suggesting that the introduction of the new law has had important effects on the behavior of individuals and their consumption of alcohol, justifying the tightening of the legislation.

JEL Classification Codes: K320- K420- R410- R480- C230

Keywords: Law Enforcement – Penalties – Deterrence – Regulatory Policies – Safety Law – Road Safety and accidents

3.1 Introduction

After heated debates that have taken place over the years on the importance of road safety, Italy has officially introduced the "Vehicular Homicide Law" (VHL) into its legislative system through the Law n. 41/2016 introduced in March 2016. The law establishes that, when, due to a traffic accident, a car driver causes the death of an individual, as a result of his own illegal behaviour, must be investigated for full-fledged homicide. The law seeks to induce citizens to a greater awareness of the responsibilities that they assume when they are driving along the roads of the country by using a tightening of legislation as a deterrent. Before the law came into force, driving while intoxicated or under the influence of drugs were not sanctioned as aggravating circumstances in the hypothesis of homicide.

The Renzi government[17], with the aim of reducing the number of tragic episodes on the streets, has decided to tighten the penalties provided by the Italian law, in light of the fact that simple prevention or education systems were ineffective in reducing tragic episodes and the number of road accidents. In practice, all citizens, when the law came into force, were called to have greater attention, greater concentration, to avoid incurring in much harsher sanctions, in the event of negligent or illegal behavior. The new legislation provides the introduction of a specific crime in the event of death on the road, following accidents or other improvised manoeuvres. In particular, the new legislation increases the years in prison, compared to the previous crime in force (simply defined manslaughter), in the case of driving offence such as driving with a blood alcohol level above the established limits or driving under the influence of drugs. In addition, the new legislation expands the number of driving offences for which detention periods are foreseen, such as speeding that causes death on the road.

Road safety is a very important issue. Every year across the streets of the most urbanized countries, several citizens lose their lives following accidents, sometimes caused by illicit conduct by drivers, especially with regard to the consumption of alcohol or drugs.

Taking Italy into consideration alone, between 2010 and 2019, 31,932 people lost their lives on the roads, while the injured amounted to more than 2 million in the same period (ISTAT, 2020). Specifically, in 2019, in Italy the "road mortality rate" (deaths per million inhabitants) was equal to 52.6, higher than the values obtained by European Union countries (EU28) equal to 48.1. Romania is the European country with the most negative values (“road mortality rate” equal to 96), while Sweden is the most virtuous (21.6). In absolute value, the highest number of road deaths was recorded in France (3,239) (ISTAT, 2020).

At the base of the consumption of alcohol by individuals there can be several causes: family problems, unemployment, social background and also other particular traits that can arise in certain territories, genetic[18]. Levitt and Porter (2011) present a methodology for measuring the risks posed by alcohol consumption for drivers that is based solely on readily available data on fatal accidents. Drivers with high level of alcohol are seven times more likely to cause a fatal accident; legally drunk drivers are 13 times more at risk than sober drivers.

In several circumstances, in the last thirty years, legislators have tried to intervene to limit these tragic episodes, using their power to generate more restrictive and harsher rules.

In addition, the increased penalties were important tools in order to study individual behaviour and their effectiveness. Traynor (2009), determined that, on average, more restrictive licensing policies for young graduates significantly reduce traffic death rates. Carpenter (2004) provided a first comprehensive analysis of the effects of the "Zero Tolerance" (ZT) law in the United States, which made it illegal to drive with any measurable amount of alcohol in the blood. Using data from 1984 to 2001, through a system called BRFSS17, he estimated with a two-way fixed effects model that the introduction of the law reduced episodic alcohol consumption among underage males by 13%. A structural and methodical evolution of Carpenter's paper was provided by Liang et al. (2008), in which, they tested the effect of the laws defined as "zero tolerance", which made it illegal to drive with any measurable amount of alcohol in the blood, considering as a sample a group of university students (under the age of 21). Using a difference-in-difference model as an estimation strategy, the authors discover that the laws of zero tolerance reduce alcohol consumption and driving among university students and that,

[17] The Behavioral Risk Factor Surveillance System (Brfss) is a state-wide health investigation system, established in 1984. It contains information on risky health behaviors, preventive clinical practices and access to health services in the 50 states, plus the District of Columbia, Puerto Rico, Guam and the Virgin Islands.

moreover, they are particularly effective in reducing the likelihood of driving after drinking for those who reported drinking outside the home. A very interesting topic in the literature is the phenomenon of recidivism episodes for car drivers, who in their past have been reported for illegal behaviour. Hansen (2015), through a RDD model, with data from the USA, determined that the aggravation of penalties leads to a reduction in illegal conduct by drivers who in the past had already received penalties for driving in a state of intoxication (both mild and moderate). Gehrsitz (2017), estimates the effects of temporary driving license suspensions on driving behavior.

Lawmakers have used very specific deterrent techniques in several countries, such as the creation of points driving license models. Bourgeon and Picard (2007), have shown that record driving licenses for society are beneficial for two reasons: they force normal drivers to drive more safely (at any time with effective mechanisms) and the removal of the license plagues more often careless drivers compared to healthy subjects. In other words, the point recording mechanisms act simultaneously as a deterrent device and as a screening and disabling device, demonstrating that non-pecuniary sanctions can be as effective as or more than monetary sanctions. De Paola, Scoppa and Falcone (2013), through a Regression Discontinuity Design, have estimated the effects of the introduction of the points license on road safety in Italy, determining that the introduction of the Penalty Points System led to a 9% reduction in road accidents and 30% of traffic fatalities.

Road accidents can be "analysed" through behavioral alterations of individuals, also using tools such as "subliminal messages". Lu et al. (2015), in a large-scale field experiment (a random sample of car owners in Tsingtao, China, divided into four groups), sent one of four cell phone text messages from police. Three groups received general messages, urging them to drive safely, warning them of the widespread use of electronic traffic monitors or describing the penalty for turning on a red light. The fourth group received personalized messages on how many traffic tickets they had received from electronic monitors. During the following month, drivers who received general messages were as likely to commit a traffic violation as drivers in the control group, while those received a personalized message committed 14% fewer traffic violations. A personalized message did not prevent subsequent violations if it simply repeated the information known to the drivers and only the new information on traffic tickets has brought significant results.

This paper brings a new contribution about the relationships existing between legislative interventions and road safety, focusing, in particular, on some behavioural dynamics that can be established in car drivers, following a tightening in penalties, in the event of unlawful conduct. The article's main objective is to verify whether the law used by the Italian government, as a deterrent to obtaining a smaller number of accidents and driving offences, had the desired effects.

To achieve this goal, we decided to implement the quasi-experiments technique, through the use of a Regression Discontinuity Design. The choice of the Regression Discontinuity Design is that this strategy is able to capture causal effects without incurring in bias deriving from temporal trends or variations in other omitted variables.

In this work we exploit data considering two different symmetrical time windows (three years) around the introduction of the new law in March 2016, controlling for a number of variables: number of patrols, price of gasoline, holidays, unemployment rate and taxes on beer. In particular, through the construction of a dataset, using information from the national police, we tested whether the introduction of the new law has led to a reduction in the number of accidents on the road (also verifying a possible reduction in the number of serious accidents) and the number of the main and most dangerous driving offenses, such as driving under the influence of drugs, driving with a too high alcohol level and the number of driving licenses withdrawn. For this reason, we present two different empirical applications: the effects of VHL, respectively, on road accidents and driving offenses.

As regards the first analysis, the dependent variables (number of accidents and number of serious accidents, causing fatalities or injured persons) show, in the period under observation, a statistically significant decreasing temporal trend. No significant effect was found with respect to the introduction of VHL.

The most interesting results are achieved in the second empirical analysis that we have conducted considering the driving offences. In fact, it is possible to observe the presence of an inverse and statistically significant relationship between the number of individuals sanctioned for having excessively high values of alcohol and the introduction of VHL (-2.628 persons sanctioned on a daily basis), suggesting that the introduction of the new law has had important effects on the behavior of individuals and their consumption of alcohol, justifying the tightening of the legislation.

This chapter is organized as follows. Section 2 briefly presents the Italian regulatory system on the subject and the innovations produced by the introduction of the Law. Section 3 shows the data and the main descriptive statistics. Section 4 presents the empirical evidence obtained and Section 5 subsequent models to verify its robustness. Section 6 sets out the conclusions.

3.2 The Italian System

The "Highway Code" is a set of laws designated to regulate the rules and behaviors to be followed by road traffic. Each country in the world defines the content, methods of execution and implementation of these rules on the basis of its internal legal system and international agreements.

The origin of this Code in Italy dates back to 1865. The current form came into force in 1993. The Code consists of 245 articles, divided into 7 sections. In particular, the fifth section defines the rules of conduct to follow in order to not endanger traffic. Article 142 governs speed limits 18.

Article 173 regulates the use of some devices while driving, stating that it is forbidden to drive using the telephone without hands-free devices or equipped with earphones. Anyone who violates this article is subject to the payment of an administrative sanction from 68.25 to 275.10 euros.

Article 186 states that driving under the influence of alcohol is prohibited. Anyone driving in a state of intoxication is punished with:

[18] For the purposes of traffic safety and the protection of human life, the maximum speed cannot exceed 130 km / h for motorways, 110 km / h for main extra-urban roads, 90 km / h for secondary extra-urban roads and for local extra-urban roads, and 50 km / h on city streets. Anyone who exceeds the maximum speed limits by no more than 10 km / h, is subject to the administrative sanction of the payment of a sum from 35 to 143 euros. Anyone who exceeds the maximum speed limits by more than 10 km / h and not more than 40 km / h. speed is subject to the administrative sanction of the payment of a sum from Euro 143 to Euro 573. Anyone who exceeds the maximum speed limits by more than 40 km / h is subject to the administrative sanction of the payment of a sum from Euro 357 to Euro 1,433.

a) a fine ranging from 500 to 2000 euros, if the blood alcohol content is greater than 0.5 and less than 0.8 grams per liter (g / l);

b) a fine ranging from 800 to 3,200 euros and imprisonment for up to six months, if the blood alcohol content is greater than 0.8 and not greater than 1.5 grams per liter (g / l);

c) a fine ranging from € 1,500 to € 6,000, arrest from three months to a year, if the blood alcohol content is greater than 1.5 grams per liter (g / l).

In addition, for a particular category of subjects, it is forbidden to drive after drinking alcoholic beverages:

a) drivers under the age of twenty-one and drivers in the first three years after obtaining the driving license;

b) drivers engaged in the activity of transporting persons;

c) drivers who carry out the transport of goods.

Article 187 states that it is forbidden to drive in conditions of physical and mental alteration related to the use of narcotic or psychotropic substances. Anyone who drives in conditions of physical and mental alteration related to the use of illegal drugs is punished with the sanctions of article 186.

The “Vehicular Homicide Law” was introduced in 2016, following a popular initiative dating back to 2011 which proposed the institution of “vehicular homicide”, a dedicated figure of crime that would impose intermediate penalties between the voluntary and culpable homicide, with arrest in the act of committing a crime and a life ban from driving vehicles (so-called "life imprisonment of the driving license").

Before the “Vehicular Homicide Law” this kind of events were prosecuted under the crime of manslaughter (penalty from 6 months to 5 years, art. 589 of the criminal code), which however provided specific aggravating circumstance for the violation of the rules of the road with sentences increased from 2 to 7 years, which could become from 3 to 10 years if the offender was found to be in a state of severe intoxication or drug abuse. The most serious offenses such as speeding and going through a red light, making U turns on bumps or curves, and driving in the wrong direction on the motorway were not punished with imprisonment.

“Vehicular Homicide Law”, provides three variants of the autonomous crime of vehicular homicide, subject to three different system of penalties.

In general, the law punishes anyone who causes, in a culpable way, the death of a person as a result of the violation of the rules governing road traffic. The penalty, in this case, is the imprisonment from two to seven years.

Different is the case in which the death of a person is caused by culpable by anyone driving a vehicle under the influence of alcohol with a blood alcohol content higher than 1.5 g / l or under the influence of illegal drugs. With reference to this case, the new law determines the imprisonment from eight to twelve years.

The last sanctioning hypothesis contemplated by the new law occurs in the case in which the death of a person is caused through fault by the driver of a vehicle who is in a state of alcoholic intoxication quantified with a blood alcohol content between 0.8 and 1.5 g / l. In this case, in fact, the penalty is the imprisonment from five to ten years.

Furthermore, it is important to point out that the new law has also tightened the penalties associated with road injures.

Also, in this case the system of penalties consists of three different levels. The general hypothesis (sanctioned with imprisonment from two to seven years in the case of homicide), is punished with imprisonment from three months to one year for serious injuries and from one to three years for very serious injuries.

The most serious cases, which are punished with imprisonment from eight to twelve years in the event of homicide, are sanctioned with imprisonment from three to five years for serious injuries and from four to seven years for very serious injuries. Finally, where the sentence of imprisonment from five to ten years for homicide is established, the new law establishes imprisonment from one year and six months to three years for serious injuries and from two to four years for very serious injuries.

Furthermore, additional changes have been introduced: compulsory arrest in the act of crime for the most serious cases (intoxication, drugs, lack of assistance); the statute of limitations for the offense has been doubled; compulsory appraisals for the collection of biological DNA samples.

3.3 Data and Descriptive Statistics

The dataset exploits different sources for the building of the variables used for the empirical evidence. It is important to clarify that in Italy there are three different "police forces", responsible for managing problems relating to road accidents: a) the national police; b) the municipal police (defined "Vigili Urbani"; c) carabinieri.

We have collected data through the information that the national police provide freely on the website http://www.poliziadistato.it/pds/stradale/archive, in which, on a daily basis, information on road accidents and driving offences is entered. which are committed within the entire national territory.

In addition, the site provides a further distinction between information relating to motorways and that relating to local roads (urban and extra-urban). Unfortunately, the other two "police forces" do not provide similar information. Using additional data, obtained from ISTAT publications, it is possible to determine that the national police, with their interventions, covered approximately 25% of the accidents recorded in Italy, in the period of time considered (the remaining part was covered by the other two law enforcement agencies).

With the awareness of the fact that we do not have the data on the totality of the accidents available, the sample of the episodes obtained can be considered representative of the totality of the incidents

Unfortunately, it is not possible to effectively use all the episodes, provided by the free datasets of the ISTAT site, since, for privacy reasons, the information relating to the exact day on which the accident occurred is omitted, indicating in a synthetic way only the quarter of the year.

The dataset used in this work takes into consideration the information relating to accidents and driving offences occurred along urban / extra-urban roads, excluding the events realized on the highways.

The data describe information on a daily basis of road accidents, traffic fatalities, injuries, different types of traffic violations carried out in the period between January 1, 2011 and December 31, 2019 (over 3,000 observations), which recorded the intervention of the national police.

In the following section, in order to make our econometric estimates, we focus on a narrower window, a symmetric 3 years time window before and after the introduction of the law.

In our analysis we use a number of control variables. Through the website of the Ministry of Economic Development[21], containing information on plants and prices applied to automotive fuels as communicated to the Ministry by the distributors' managers in implementation of art. 51 L. 99/2009, we have reported the national average of the cost of fuel, using weekly information, in order to control for traffic intensity (De Paola, Scoppa and Falcone, 2013). Also, to isolate the effect of a possible increase in the price on the consumption of the alcohol, we have obtained information on the excise duties paid on beers, directly from the annual reports published by AssoBirra, which represents the heart of the Italian beer supply chain and brings together the main companies that produce and market beer and malt in Italy, considering the value of one hectolitre of beer with 12 degree Plato as a reference point. We did not use ISTAT data as it is not possible to disaggregate the information on alcohol prices from those on tobacco.

Finally, in order to consider the economic situation of the country, we added information about the unemployment rate to the estimates, through the monthly tables published by ISTAT.

In addition, through the national police site, we also collected information on the number of patrols employed.

Table 3.1 shows the descriptive statistics for the main variables used considering the entire sample. The average daily number of road accidents was 76.745. Among these, the average number of accidents with injuries was 42.716 (55.66%) and the average number of fatal accidents was 1.518 (1.98%).

The average daily number of people sanctioned for drunk driving is equal to 39.702, while an average of 2.803 people reported sanctions for driving under the influence of drugs.

Table 3.1. Descriptive Statistics

Variable	Obs	Mean	Std.Dev.	Min	Max
Police Patrols	3287	685.249	102.68	102	1147
Road Accidents	3287	76.745	17.05	29	153
Fatalities	3287	1.518	1.311	0	10
Injuries	3287	42.716	11.533	13	87
Alcool	3287	39.702	32.399	4	260
Unemployment Rate	3287	11.045	1.314	7.9	13.3
Beer Tax	3287	2.764	0.31	2.33	3.04
Gasoline	3287	1.609	0.117	1.361	1.89
Speed	3287	473.201	380.46	3	4662
Phone	3287	64.423	35.053	3	326
Drug	3287	2.803	2.327	0	20
Driving License	3287	130.21	42.431	30	771

Daily data: 1st January 2011 to 31st December 2019.
Sources Italian national police, Ministry of Economic Development, *ISTAT* Italian statistical office, AssoBirra.

Table 3.2 provides information about the descriptive statistics of the main variables used in the analyses separately for the period before and the period following the introduction of the law covered by the paper. Almost all the variables reported significant downward variations considering the two distinct periods.

The number of accidents decreased by 22.038 (passing from 87.728 to 65.690), as well as consequently the number of traffic injuries of 14.244 (passing from 50.011 to 35.766) and the number of fatalities by 0.544 (passing from 1.799 to 1.255).

In absolute terms, the highest variation was recorded for the Speed variable. On average, the number of people sanctioned for speeding decreased by 174.333 (from 589.751 to 415.417). Strong reduction is recorded also for the number of motorists sanctioned as results with alcohol values higher than those allowed (passing from 45.790 to 34.323).

The only driving offenses that increased in the two periods under study is the number of people sanctioned for using mobile phones while driving (+14.813).

As for the control variables, there was a drastic reduction in the number of patrols employed by the police and no significant difference between the two periods considering the unemployment rate.

Table 3.2: Descriptive Statistics before and after the introduction of VHL

	Before The Law (Mean)	After the Law (Mean)	Differences
Road Accidents	87.728	65.69	-22.038*** (0.516)
Police Patrols	756.374	607.467	-148.908*** (3.004)
Fatalities	1.799	1.255	-0.544*** (0.049)
Injuries	50.011	35.766	-14.244*** (0.347)
Alcohol	45.790	34.323	-11.466*** (1.065)
Unemployment Rate (monthly)	10.869	10.824	-0.044 (0.064)
Gasoline Price	1.703	1.545	-0.158*** (0.002)
Speed	589.751	415.417	-174.333*** (14.568)
Phone	58.972	73.785	+14.813*** (1.321)
Drugs	2.928	2.801	-0.127** (0.089)
Driving License	141.2614	121.744	-19.517*** (1.489)

Daily data: 1st January 2011 to 31st December 2019. The symbols *** indicate that the difference is statistically significant at the 1 % level

Sources Italian national police, Ministry of Economic Development, *ISTAT* Italian statistical office, AssoBirra.

3.4 Empirical Analysis and Results

The main objective of the paper is to try to understand if sharp increase of penalties, culminating with the creation of "Vehicular Homicide Law", has led to a reduction in the number of road accidents and in the number of driving offences.

To achieve this goal, we use a Regression Discontinuity Design (RDD), introduced for the first time in the literature by Thistlewaite and Campbell in 1960[21], exploiting the creation of a "threshold" value that allows the division of the sample into two groups: treatment and control.

To apply this model correctly, we had to consider a series of properties: a) perfect assignment to the two groups through the threshold value; b) non-existence of other factors that could create the discontinuity near the threshold; c) the treatment is the same for all those who are exposed to it.

Specifically, we define the dummy variable VHL_i equal to one for all the observations recorded after the entry into force of the law at t_0 = 25th March 2016 and 0 otherwise:

$$VHL_i = \begin{cases} 1 & if\ t \geq 25th\ March\ 2016 \\ 0 & if\ t < 25th\ March\ 2016 \end{cases}$$

In particular, we used a Sharp Regression Discontinuity Design, where the the treatment is a deterministic and discontinuous function of time (Angrist et al., 2009). Specifically, the observations recorded in the days prior to March 25, 2016 are considered not subjected to treatment (the law had not yet entered into force) and consequently, the information obtained from the data referring to the following days are considered subjected to treatment. The goal of these econometric tools is to compare the outcomes variables Y among the subjects located around the threshold. Clearly, the number of road accidents and the number driving offences could be correlated to a number of other factors: road maintenance and investments in infrastructures, technological progress of tools for detecting alcohol levels, types of interventions by law enforcement, technological progress in machine safety tools. The fundamental assumption of the model is that these temporal trends are smooth around the threshold.

We model the number of per day road accidents, fatalities, injuries and driving offences, using the following model:

$$Y_t = \alpha + \beta VHL_t + \gamma f(\text{time}) + \delta X_t + \varepsilon_t \quad (1)$$

Where Y_t is our outcome variable (respectively accidents, fatalities, injuries and driving offences) for day t, the coefficient β represents the effect of the treatment, f(time) is a function of time and X is a vector of control variables and ε is the error term.

The entire reference time range is nine years (3287 daily data), following what is widely used in the literature, we used the observations with a Δ distance on both sides of the t_0 threshold: [t_0-Δ, t_0+Δ]. For this reason, we have created two different symmetrical windows respectively to the right and left of the reference threshold (3 years symmetric window). As a robustness check, in Sect. 5.1, we estimate our model focusing on alternative time windows.

3.4.1 The effects of "Vehicular Homicide Law" on road accidents

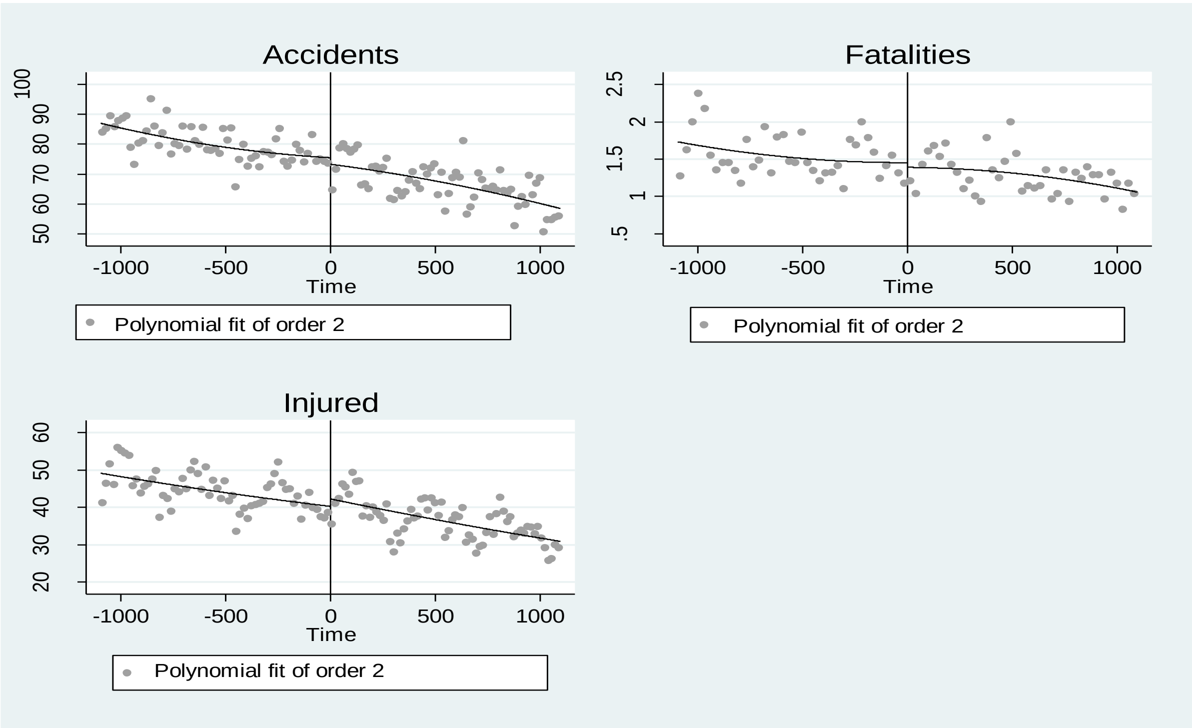

Figure 3.1: RD plot for the impact of the Vehicular Homicide Law on, respectively, Accidents, Fatalities and Accidents with Injured

Figure 3.1 shows the number of accidents, the number of fatalities and the number of injured, as time varies (using a second order polynomial). In all cases there is a decreasing trend of the phenomena, without observing a marked discontinuity near the threshold.

Tables 3.3, 3.4 and 3.5 show OLS estimates of Eq.1 for, respectively, accidents, fatalities and injuries.

Considering Table 3.3, it emerges that, at the same time as the introduction of the new law, there was a reduction in the number of daily accidents of 1.25 (column 3). This evidence, however, does not appear to be statistically significant.

The "time" variable plays a fundamental role in all specifications, which confirms a decreasing trend in the number of accidents during the period under analysis.

Furthermore, there is a significant relationship between the price charged for fuel in Italy and the total number of accidents. As the price increases, the number of accidents decreases (-6.283).

Table 3.4 confirms the presence of a negative but not statistically significant relationship between the number of fatalities and the introduction of the VHL with a coefficient equal to -0.05 (column 3).

In Table 3.5 we can draw conclusions similar to those obtained from Table 3.3, confirming a correlation between the number of road accidents and the number of injuries, also as regards the control variables used. Specifically, the introduction of VHL led to a reduction in the number of injured by 0.653 (column 3).

The evidence obtained in these tables lead to the assertion that the law introduced in Italy has not caused any causal effects on the number of daily road accidents, nor on the number of more serious road accidents.

Therefore, we can assert that any behavioral changes induced by the introduction of the new law have not overall had a significant effect on the number of road accidents and the number of victims.

Table 3.3: Regressions for the impact of the Vehicular Homicide Law on the number of Accidents in Italy (2013-2019)

	Accidents	Accidents	Accidents
VHL	-0.191	-0.138	-1.250
	(1.070)	(1.070)	(1.094)
Time	-0.012***	-0.013***	-0.011***
	(0.000)	(0.001)	(0.001)
Police Patrols		0.004	0.026***
		(0.004)	(0.006)
Holidays		-1.485*	-2.336
		(0.782)	(1.716)
Unemp. Rate		-0.580	-1.192
		(0.793)	(0.788)
Gasoline		-2.835	-6.283**
		(2.902)	(2.923)
Days dummies			YES
Monthly dummies			YES
Constant	73.365***	82.552***	77.335***
	(0.205)	(10.668)	(10.959)
Observations	2193	2193	2193
Adjusted R^2	0.269	0.273	0.323

The table reports RD estimates. All specifications are for 25th March 2013– 25th March 2019, considering also dummies variables for month of the year and day of the week. Standard errors (reported in parentheses) are corrected for heteroskedasticity. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10% level.

Table 3.4: Regressions for the impact of the Vehicular Homicide Law on the number of Fatalities in Italy (2013-2019)

	Fatalities	Fatalities	Fatalities
VHL	-0.035	0.021	-0.050
	(0.102)	(0.101)	(0.104)
Time	-0.000**	-0.000***	-0.000
	(0.000)	(0.000)	(0.000)
Police Patrols		-0.002***	-0.000
		(0.000)	(0.001)
Holidays		0.291***	0.015
		(0.082)	(0.155)
Unemp. Rate		0.020	0.021
		(0.080)	(0.081)
Gasoline		-0.035	-0.229
		(0.264)	(0.268)
Days dummies			YES
Monthly dummies			YES
Constant	1.372***	2.477**	1.995*
	(0.116)	(1.043)	(1.089)
Observations	2193	2193	2193
Adjusted R^2	0.018	0.043	0.062

The table reports RD estimates. All specifications are for 25th March 2013– 25th March 2019, considering also dummies variables for month of the year and day of the week. Standard errors (reported in parentheses) are corrected for heteroskedasticity. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10% level.

Table 3.5: Regressions for the impact of the Vehicular Homicide Law on the number of Injured in Italy (2013-2019)

	Injured	Injured	Injured
VHL	-2.026**	-1.893***	-0.653
	(0.714)	(0.706)	(0.679)
Time	-0.009***	-0.012***	-0.006***
	(0.000)	(0.001)	(0.001)
Police Patrols		-0.019***	0.015***
		(0.003)	(0.004)
Holidays		0.252	1.340
		(0.575)	(1.200)
Unemp. Rate		-0.760	-0.626
		(0.552)	(0.520)
Gasoline		2.340	-4.251**
		(1.872)	(1.769)
Days dummies			YES
Monthly dummies			YES
Constant	36.689***	57.585***	39.975***
	(0.735)	(7.114)	(6.945)
Observations	2193	2193	2193
Adjusted R^2	0.261	0.281	0.402

The table reports RD estimates. All specifications are for 25th March 2013– 25th March 2019, considering also dummies variables for month of the year and day of the week. Standard errors (reported in parentheses) are corrected for heteroskedasticity. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10% level.

3.4.2 Effects of "Vehicular Homicide Law" on driving offences

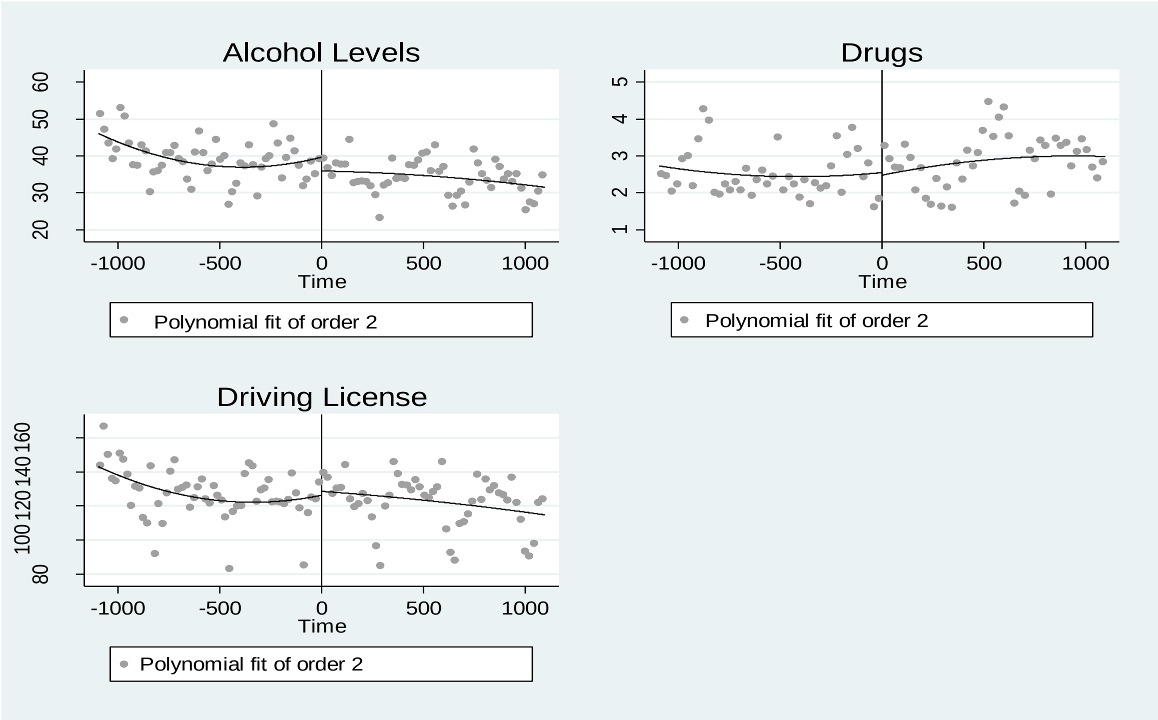

Figure 3.2: RD plot for the impact of the New Law on Driving Offences

Figure 3.2 analyses the trend in the number of motorists sanctioned for driving under the influence of alcohol, for driving under drug use and the number of licenses withdrawn, as the "time" variable changes. We can observe the presence of a discontinuity near the threshold as regards the number of drunk driving.

Tables 3.6, 3.7, 3.8 show OLS estimates of Eq.1 for, respectively, the number of drivers sanctioned for drunk driving, the number of drivers sanctioned for driving under illegal drug and the number of driving licenses withdrawn.

Focusing on Table 3.6, we can observe the most interesting results.

In fact, from column 3 it emerges that, by controlling for a series of variables including temporal ones, the introduction of VHL has led to a reduction in the number of people sanctioned for drunk driving of 2.628, that is, a reduction of about 6%. This coefficient is statistically significant at the 5% level.

As regards the other variables that make up the model, predictably, the variable Holidays has a strong impact on the dependent variable: in these periods of the year, the number of drivers sanctioned for drunk driving were very high (+22.573). Gasoline price and beer tax have a negative impact on the dependent variable, contrary to the number of patrols which, predictably, positively affect the number of people sanctioned for driving under the influence.

As regards the last two types of driving offences (Tables 3.7 and 3.8), the results obtained do not show statistical significance in each specifications, indicating that the law introduced to the Italian government has not had any impact on these two phenomena.

Table 3.6: Regressions for the impact of the Vehicular Homicide Law on drunk driving in Italy (2013-2019)

	Under Influence of Alcohol	Under Influence of Alcohol	Under Influence of Alcohol
VHL	-0.423	-0.785	-2.628**
	(2.519)	(2.066)	(1.238)
Time	-0.005**	-0.013***	-0.002
	(0.002)	(0.003)	(0.002)
Police Patrols		-0.049***	0.022**
		(0.008)	(0.009)
Holidays		45.701***	22.753***
		(1.630)	(3.029)
Unemp. Rate		-2.926*	-2.348**
		(1.635)	(0.969)
Gasoline		4.044	-9.058**
		(7.641)	(4.607)
Beer Tax		3.277	-6.829**
		(5.183)	(3.100)
Days dummies			YES
Monthly dummies			YES
Constant	36.549***	80.212***	104.457***
	(1.399)	(26.497)	(15.273)
Observations	2193	2193	2193
Adjusted R^2	0.009	0.409	0.806

The table reports RD estimates. All specifications are for 25th March 2013– 25th March 2019, considering also dummies variables for month of the year and day of the week. Standard errors (reported in parentheses) are corrected for heteroskedasticity. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10% level.

Table 3.7: Regressions for the impact of the Vehicular Homicide Law on people sanctioned for driving under drug in Italy (2013-2019)

	Under Influence of Drugs	Under Influence of Drugs	Under Influence of Drugs
VHL	0.175	0.155	0.022
	(0.198)	(0.210)	(0.226)
Time	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)
Police Patrols		-0.000	0.002**
		(0.001)	(0.001)
Holidays		0.590***	0.028
		(0.180)	(0.381)
Unemp. Rate		-0.376**	-0.474***
		(0.174)	(0.182)
Gasoline		0.317	-0.373
		(0.745)	(0.779)
Beer Tax		-0.216	-0.386
		(0.508)	(0.526)
Days dummies			YES
Monthly dummies			YES
Constant	2.594***	7.287***	8.535***
	(0.110)	(2.380)	(2.359)
Observations	2193	2193	2193
Adjusted R^2	0.005	0.017	0.054

The table reports RD estimates. All specifications are for 25th March 2013– 25th March 2019, considering also dummies variables for month of the year and day of the week. Standard errors (reported in parentheses) are corrected for heteroskedasticity. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10% level.

Table 3.8: Regressions for the impact of the Vehicular Homicide Law on Driving License Withdrawn in Italy (2013-2019)

	Driving License withdrawn	Driving License withdrawn	Driving License withdrawn
VHL	10.296***	13.873***	1.393
	(3.145)	(3.161)	(2.923)
Time	-0.014***	-0.031***	-0.003
	(0.003)	(0.005)	(0.005)
Police Patrols		0.007	0.092***
		(0.012)	(0.017)
Holidays		40.035***	14.240***
		(2.965)	(4.458)
Unemp. Rate		-12.251***	-6.111***
		(2.314)	(2.089)

Gasoline		37.319***	-8.875
		(11.803)	(11.997)
Beer Tax		21.100***	-7.430
		(7.622)	(7.124)
Days dummies			YES
Monthly dummies			YES
Constant	119.772***	129.704***	180.141***
	(1.672)	(38.911)	(32.218)
Observations	2193	2193	2193
Adjusted R^2	0.015	0.154	0.437

The table reports RD estimates. All specifications are for 25th March 2013– 25th March 2019, considering also dummies variables for month of the year and day of the week. Standard errors (reported in parentheses) are corrected for heteroskedasticity. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10% level.

3.5. Robustness Check

3.5.1 Poisson Estimator

In light of the fact that our dependent variables take non-negative discrete values (De Paola, Scoppa and Falcone, 2013), it is appropriate to implement the reference model using a Poisson estimator, in order to provide further robustness to the results obtained in the previous section.

Replicating the specifications of Table 3.6 in Table 3.9, considering the phenomenon of drunk driving, with this new estimator we have obtained statistically significant results, consistent with the OLS estimates of Table 3.6. In fact, from column 3, it emerges that with the entry in force of the new law, all other conditions being equal, the number of people punished for drunk driving is about 6.9% lower than the previous period.

Table 3.9: RDD estimates of VHL Effects on drunk driving in Italy (Poisson Estimates, 2013-2019)

	Under Influence of Alcohol	Under Influence of Alcohol	Under Influence of Alcohol
VHL	-0.011	-0.032	-0.069**
	(0.069)	(0.058)	(0.035)
Time	-0.000**	-0.000***	-0.000
	(0.000)	(0.000)	(0.000)
Police Patrols		-0.001***	0.001**
		(0.000)	(0.000)
Holidays		0.905***	0.612***
		(0.031)	(0.082)
Unemployment Rate		-0.082*	-0.061***
		(0.044)	(0.024)
Beer Tax		0.113	-0.158**
		(0.137)	(0.079)
Gasoline		0.108	-0.236*
		(0.210)	(0.124)
Day dummies			YES
Monthly dummies			YES
Constant	3.596***	4.707***	4.888***
	(0.038)	(0.692)	(0.379)
Observations	2193	2193	2193

The table reports RD estimates considering Poisson distribution. All specifications are for 25th March 2013– 25th March 2019, while in columns (3) and (4) I report indicator variables for month of the year and day of the week. Standard errors (reported in parentheses) are corrected for heteroskedasticity. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10% level.

In order to provide more information on the previous evidence, Table 3.10 shows the specifications obtained through the Poisson estimator, replicating what was done with the OLS models for each dependent variable, obtaining similar results with the previous estimates contained in section 3.4.2.

Table 3.10: RDD estimates of VHL Effects on the others dependent variables in Italy (Poisson Estimates, 2013-2019)

	Accidents	Fatalities	Injured	Drugs	D.L. withdrawn
VHL	-0.018	-0.036	-0.018	0.001	0.009
	(0.015)	(0.073)	(0.016)	(0.088)	(0.024)
Time	0.000***	-0.000	0.000***	0.001	0.005
	(0.000)	(0.000)	(0.000)	(0.071)	(0.022)
Police Patrols	0.000***	-0.000	0.000***	0.001**	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Holidays	-0.032	0.012	0.035	0.000	0.113***
	(0.024)	(0.114)	(0.030)	(0.152)	(0.036)
Unem_Rate	-0.010	0.019	-0.007	-0.163**	-0.048***
	(0.010)	(0.052)	(0.012)	(0.066)	(0.016)
Gasoline	-0.126***	-0.226	-0.170***	-0.231	-0.057
	(0.040)	(0.189)	(0.044)	(0.297)	(0.098)
Beer Tax				-0.211	-0.398***
				(0.203)	(0.041)
Day dummies	YES	YES	YES	YES	YES
Monthly dummies	YES	YES	YES	YES	YES
Constant	4.325***	0.729	3.649***	3.299***	5.144***
	(0.140)	(0.718)	(0.160)	(0.922)	(0.254)
Observations	2193	2193	2193	2193	2193

The table reports RD estimates considering Poisson distribution. All specifications are for 25th March 2013– 25th March 2019 and we report indicator variables for month of the year and day of the week. Standard errors (reported in parentheses) are corrected for heteroskedasticity. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10% level.

3.5.2 Further specifications with different functions of time

A potential threat to the internal validity of the RDD is the incorrect specification of the function that relates Y to X. For this reason, in light of what is highlighted in the previous tables, considering the phenomenon of "drunk driving", it is appropriate to present models with different functions of the "forcing variable", inserting interactions and regressions with polynomial forms.

Contrary to a common practice in the literature over the years, Gelman and Imbens (2019) argue for not using high-order polynomials in the Regression Discontinuity Design. For this reason, in Table 3.11 we only use the Time^2 variable, to verify the existence of a possible quadratic relationship between the dependent variable and time.

Furthermore, we use an interaction term between Time and VHL to model different functional forms on the two sides of the cut-off.

In Column 4, we used polynomials of Time and interaction term together.

In each estimate shown in Table 3.11, the introduction of the new law has led to a reduction in the number of individuals sanctioned with a blood alcohol level higher than permitted, consistent with what is shown in the previous section.

Table 3.11: Regressions with polynomial form and interaction terms for Drunk Driving in Italy (2013-2019)

	Under Influence of Alcohol	Under Influence of Alcohol	Under Influence of Alcohol	Under Influence of Alcohol
VHL	-2.628**	-2.801**	-2.768**	-2.790**
	(1.238)	(1.329)	(1.305)	(1.327)
Time	-0.002	-0.002	-0.000	-0.001
	(0.002)	(0.002)	(0.004)	(0.006)
Time^2		-0.000		-0.000
		(0.000)		(0.000)
Time*VHL			-0.003	-0.002
			(0.005)	(0.010)
Constant	61.138***	78.902***	108.837***	109.218***
	(15.097)	(21.408)	(17.390)	(23.279)
Observations	2193	2193	2193	2193
Adjusted R^2	0.806	0.809	0.809	0.809

The table reports RD estimates. All specifications are for 25th March 2013– 25th March 2019 and include indicator variables for month of the year, day of the week, police patrols, beer tax, gasoline price, unemployment rate, dummies for holiday. Standard errors (reported in parentheses) are corrected for heteroscedasticity. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5 and 10% level.

Table 3.12 presents the regressions using the polynomial forms of the variable "time" and the interaction term between time and VHL, considering the other dependent variables. Further, through these specifications, the results demonstrated in the previous sections are confirmed.

Table 3.12: Regressions with polynomial form and interaction terms for the others dependent variables in Italy (2013-2019)

	Accidents	Fatalities	Injured	Drugs	D.L. withdrawn
Time	-0.014**	-0.000	-0.002	-0.003**	-0.000
	(0.006)	(0.001)	(0.004)	(0.001)	(0.016)
Time^2	-0.000	-0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Time*VHL	0.005	0.000	-0.008	0.005**	-0.004
	(0.010)	(0.001)	(0.006)	(0.002)	(0.026)
VHL	-1.026	-0.116	-0.541	0.073	1.137
	(1.343)	(0.116)	(0.815)	(0.244)	(3.323)
Constant	76.416***	3.434*	30.934**	9.877**	187.584***
	(22.615)	(1.993)	(13.482)	(3.980)	(63.151)
Observations	2193	2193	2193	2193	2193
Adjusted R^2	0.322	0.062	0.401	0.057	0.437

The table reports RD estimates. All specifications are for 25th March 2013– 25th March 2019 and include indicator variables for month of the year, day of the week, police patrols, beer tax, gasoline price, unemployment rate, dummies for holiday. Standard errors (reported in parentheses) are corrected for heteroscedasticity. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5 and 10% level.

3.5.3 Different time windows

Taking into consideration what Imbens and Lemieux (2008) demonstrated, a further robustness check consists in reducing the size of the time windows used for empirical analyses. This comparison allows us to estimate the effects in such a way as not to incur an error due to the incorrect specification of the time function. We focus on three different time windows: the first which includes two years before and two years after t_0, the second which considers a year and a half before and a year and a half after t_0, the third focused on a window of one year.

Table 3.13 shows the results with the alternative time windows. In each specification, the introduction of the new law has led to a reduction in the number of people sanctioned for drunk driving. It is important to emphasize that, as the observations in the time windows used for the analysis are reduced, the effect of the law appears to be, in terms of magnitude, smaller. Furthermore, in the third window, we see an absence of statistical significance. However, it is necessary to underline the importance of seasonal effects on our variable (the number of drivers punished for drunk driving is much higher at certain times of the year: summer or Christmas holidays), the results obtained could be biased towards the zero (De Paola and Scoppa, 2012).

Table 3.13: RD estimates with alternative time windows for Drunk Driving in Italy

Time window	Under Influence of Alcohol
25/03/2014-25/03/2018	-4.749**
	(1.638)
Observations	1462
25/09/2014-25/09/2017	-3.777**
	(1.887)
Observations	1091
25/03/2015-25/03/2017	-2.338
	(3.232)
Observations	730

The table reports RD estimates of New Law. All specifications include indicator variables for month of the year, day of the week, police patrols, beer tax, gasoline price, unemployment rate, dummies for holiday. Standard errors (reported in parentheses) are corrected for heteroscedasticity. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5 and 10% level.

3.5.4 A nonparametric approach

Nonparametric local polynomial estimators represent important tools in RD literature, determining an important robustness check for the validity of the estimates presented.

It is kind of regression study in which the predictor does not take a determined form but is made built on evidence provided from the data. Specifically, instead of estimating the parameters of a specific functional form, the functional form itself is estimated (Jacob et al., 2012). This estimation strategy only considers observations in a symmetric range around the threshold (the bandwidth), using weighted polynomial regressions, typically of order 1 or 2, with weights calculated by applying a kernel function (Cataneo et al., 2015).

The confidence intervals are constructed using what was created by Calonico, Cataneo and Titiunik (2015). In this paper we used the suite "rdrobust" in STATA, considered a second order polynomial regression and different ranges of observations.

In Table 3.14, the empirical evidence provided refers to four different time windows around the threshold (bandwidth) expressed in days: 270, 365, 545, 730, consequently considering

4 different values of observations. While maintaining negative coefficient values, the results obtained do not show significant effects of the variable of interest on the number of drivers sanctioned as they are under the influence. It is observed that the lower is the value of the bandwidth chosen, the lower is the value of VHL. The explanation for this type of result lies in the fact that, since time is our forcing variable, the results obtained suffer from seasonal effects, following what was observed in the previous section (3.5.3).

Table 3.14: RD estimates with a nonparametric local polynomial model for Drunk Driving in Italy

	U.I. of Alcohol	U.I. of Alcohol	U.I. of Alcohol	U.I. of Alcohol
VHL	-0.970	-1.330	-3.681	-3.901
	(2.785)	(2.256)	(2.512)	(2.660)
Bandwidth	270	365	545	730
Observations	540	730	1090	1460

The table reports RD estimates of New Law using a nonparametric approach. All specifications include robust confidence interval estimators for average treatment effects at the cutoff in sharp RD following Calonico, Cataneo and Titiniuk (2014). Each column has different bandwidth values expressed in days, respectively 270 (nine months around the threshold), 365 (one year around the threshold), 545 (one year and a half around the threshold), 730 (two years around the threshold). The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5 and 10% level.

3.6 Concluding Remarks

This work analyzes whether the increase in deterrence levels caused by the introduction "Vehicular Homicide Law" in Italy in March 2016 has led to a decrease in the number of road accidents (more or less serious) and in the number of driving offences.

After collecting data from different sources (mainly from the official website of the National Police), we used a Regression Discontinuity Design to compare road accidents and driving offences before and after the introduction of VHL, checking for several variables: number of police patrols, holiday days, unemployment rate, gasoline price, seasonal effects.

The introduction of VHL did not significantly influence the total number of accidents or the number of serious accidents (fatalities and injured): we observe a general decreasing trend of these variables.

On the other hand, we have found significant results through the analysis of driving offences. In particular, we observed the existence of a negative and statistically significant relationship between VHL and people sanctioned because in a state of intoxication.

Specifically, the introduction of the “Vehicular Homicide Law” led, on average, to a reduction of 2.628 people sanctioned for drunk driving.

Furthermore, in general, it emerges that on driving offenses, the unemployment rate has a significant negative impact. As the unemployment rate increases, driving offenses (alcohol abuse, drug abuse and the number of driving licenses withdrawn) are significantly reduced.

In addition, we proceeded using a Poisson estimator, since our dependent variable takes non-negative discrete values: the results obtained are robust with those obtained in the OLS model. These findings are robust to different polynomial time trends and to the use of different time windows (except smaller ones, probably due to seasonal effects).

We can conclude that the law has had partial but significant results on one of the main and most dangerous driving offenses: drunk driving.

In addition, this article shows that it is possible to influence individual behaviours through legislative interventions. For fear of harsher sanctions, individuals tend to engage in more careful behaviours.

3.7 References

Angrist J., Pischke J. (2009) "Mostly harmless econometrics", Princeton University Press, London;

Annual Report Associazione dei Birrai e dei Maltatori 2013-2019;

Annual Report ISTAT Incidenti Stradali periodo di riferimento 2019;

Becker G. S. (1968) "Crime and punishment: an economic approach", Journal of Political Economy, 76, 160-217;

Benson B. L., Kim I. and Rasmussen D. W. (1994) "Estimating deterrence effects: a public choice prospective on the economics of crime literature", Southern Economic Journal, 61 (July), 161-168;

Bourgeon J., Picard P. (2007) "Point-record driving licence and road safety: an economic approach", Journal of Public Economy, 91:235–258;

Benson B.L., Mast B.D., Rasmussen D.W. (2000) "Can police deter drunk driving?", Applied Economics, 32:3, 357-366;

Carpenter C. (2004) "How Do Zero Tolerance Drunk Driving Laws Work?", Journal of Health Economics, 23 (1): 61–83;

Carpenter C. and Dobkin C. (2009) "The Effect of Alcohol Consumption on Mortality: Regression Discontinuity Evidence from the Minimum Drinking Age", American Economic Journal: Applied Economics, (1): 164–82;

De Paola M., Scoppa V., Falcone M. (2010) "The deterrent effects of the penalty point system for driving offences: a regression discontinuity approach" Working Paper n. 4-2010 Dipartimento di Economia e Statistica Università della Calabria, Febbraio;

Gehrsitz M. (2017) "Speeding, Punishment and Recidivism. Evidence from a Regression Discontinuity Design", Journal of Law and Economics, 60(3);

Imbens G. W. and Lemieux T. (2008) "Regression Discontinuity Designs: A Guide to Practice." Journal of Econometrics, 142 (2): 615–35;

Liang L. and Huang J. (2008) "Go out or stay in? The effect of zero tolerance laws on alcohol use and drinking and driving patterns among college students", Health Economics, 17(11);

Lu F., Zhang J., Perloff J.M. (2015), “General and specific information in deterring traffic violations: Evidence from a randomized experiment”, Journal of Economic Behavior & Organization, 123: 97-107;

Jacob R., Zhu P., Somers M., Bloom H. (2012). “ A Pratical Guide to Regression Discontinuity”, MDRC Publications.

Hansen B. (2015) “Punishment and Deterrence: Evidence from Drunk Driving”, American Economic Review, 105(4): 1581-1617;

Levitt, Steven D., and Jack Porter. 2001. “How Dangerous Are Drinking Drivers?” Journal of Political Economy 109 (6): 1198–1237;

McCrary J. (2008) “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.” Journal of Econometrics, 142 (2): 698–714;

Thistlethwaite D. L. and Campbell D.T. (1960) “Regression-discontinuity analysis: An alternative to the ex post facto experiment.” Journal of Educational Psychology 51 (6): 309–17;

Traynor T. (2009) “The impact of state level behavioral regulations on traffic fatality rates”, Journal of Safety Research, 40:421–426;

Yu J. (2000) “Punishment and alcohol problems: Recidivism among drinking-driving offenders.” Journal of Criminal Justice 28 (4): 261–70;

Yu J., Evans P.G., Clark L.P. (2006) “Alcohol addiction and perceived sanction risks: Deterring drinking drivers”, Journal of Criminal Justice, 34: 165-174.