

# Betting on poverty

## A comparative study between Barcelona and Madrid

*Gerard Adell, Albert Canadell, Marc Cánovas, Daniel Soldevila*

*Director of the project: Juan José Ganuza*

*Code: GJE01*

Treball de fi de grau en Administració i Direcció d'Empreses (UPF). Curs 2018-2019

### ABSTRACT

The debate about increasing gambling is in the spotlight of the whole society. At the same time, several drives join together to claim for a tighter regulation in view of the raising amount of compulsive gamblers. Our article seeks to reveal the existence of a significant impact of household income on the location decision of betting houses within neighborhoods and whether or not spatial autocorrelation between them plays a big role. Enclosing Barcelona and Madrid as study areas, our analysis shows an inverse relationship between the number of betting houses per neighborhood and the economic status of its residents, but levels of significance differ across cities. Thus, this study aims to understand the contrast of results among them, identifying regulatory framework as a potential distinguishing element.

## INTRODUCTION

### *An overview of the global gambling background*

The pattern of a new trend of compulsive gambling is growing exponentially at the rate of the gaming industry's profit (Toca, 2017). We are in front of a fast-growing industry, both in terms of financial figures and of social repercussion issues, which is exposed to a ceaseless and dynamic change by the hand of new technologies.

The sector embraces both physical and virtual worlds. Because of its huge revenues and high tax rates it is exposed to, its fiscal contribution is tremendous. However, it also triggers more than 50,000 suicides annually (Jones, 2018). What kind of weapon are bets for society?

The evolution trend of the sports betting business across the world does not seem to stop, accounting a revenue of \$183 billion in 2015 (Fenez, 2016) and foresees nothing more than \$525 billion by 2023 (Research and Markets, 2017). At national level, in Spain, the whole gambling industry moves €35,000 million annually, the equivalent of 3.5% of the GDP and 80,000 direct jobs depend on it (Toca, 2017).

Despite the rise of new technological tools and the huge exposure of international online operators, the Spanish market is still ruled by on-site betting houses, whom absorbed more than 75% of the sector's revenues and 95% of margins in 2015 (Gómez et al., 2016). And it is in this field where premises' strategic location becomes highly important.

In this way, betting houses look indiscriminately for being settled as close to the main consumer target as possible. Target's profile has evolved in the last years in parallel with the sector. According to specialized sources, the average age of the gambler prototype has dropped considerably in recent years. Following a study of 2012, the general picture of participation in sports betting markets in Canada, Spain and the UK is that sports bettors tend to be younger males with relatively high income (Humphreys and Pérez, 2012). Nevertheless, the amount and range of bettors they reach is so wide that is hard to define just a single profile.

All the before mentioned makes up a brief reason why this issue is so current in today's society. And it leads us to ask: How are betting houses strategically locating to take advantage of society's weaknesses?

### *The Spanish market context*

Spain's landscape has played a fundamental role in the development of betting, not only due to the its influence in sports across the world, but also because of its market performance and singular legal framework, which will be analyzed in the following lines.

### *Market mechanisms*

The Spanish gambling market is divided between public and private operators, which combined result in 54 licensed operators as registered in the DGOJ (Dirección General de Ordenación del Juego, 2018) which is the management body of the Ministry of Finance.

Considering general figures, the Spanish gambling market has undergone a process similar to most of the economics sectors in Spain regarding the severe impact of the financial crisis. Its Gross Gambling Revenue (GGR), which is the difference between bet amounts less prizes, had consistently increased until its 2008 peak when the crisis dropped its total amount by 26,7% within six years.

Starting in 2015, a consistent increasing trend demonstrates the sector is impressively recovering although pre-crisis levels haven't yet been reached. An explanation of the slump is the fact gambling is an expendable service for almost all users, except for those who develop addiction towards it.

### *Regulatory framework*

The growth of the betting market in Spain triggered the need of establishing a regulatory framework with the main objective settled in smoothing the apparent legal impunity derived from these kind of markets – tax evasion, fraud, addictive pathologies, etc. –.

Since the betting sector is immersed in a permanent process of technological innovation, most of the measures are aimed at regulating Spain's online betting landscape. This is the precedent of the Spanish Gambling Act of 2011.

Quoting BOE's bulletin, "New Law 13/2011 on Gaming («Gaming Law»), provides a legal certainty and a general regulatory framework to online state-wide gaming activities in Spain. The Gaming Law divides the gaming market into two separate segments: (i) the reserved market, which may only be carried out by the entities designated in the Gaming Law (State-wide lottery and lotto games companies), and (ii) the non-reserved market, where the operation of games other than lottery and lotto is open to competition although prior administrative license is required."

The regulatory law has caused the appearance of huge gambling multinationals in Spain, and thus the reduction of the state monopolistic profits collected from its own company, SELAE (*Sociedad Estatal Loterías y Apuestas del Estado S.A.*), as its profits are eroded consequence of increased competition.

However, gambling is still a considerable income source of state public purse thanks to corporate taxes and special game taxes applied, although some of this important multinationals locate its tax domicile in more tax-lax states such as Malta, Cyprus or Gibraltar (Fernández, 2018).

Moreover, even though the legal framework embraces several competences that do not really constitute the game as an activity (e.g. advertising and promotion), most of the Gaming Law's measures are addressed to the non-reserved market, specially regulating the activity when it is carried out through electronic, computer, telematics and interactive channels (González-Espejo and López, 2018). Thus, the online betting market.

Therefore, while the Spanish Government is in charge of managing the online betting, Autonomous Communities are the ones who bear the competencies regarding the regulation of on-site betting activities. And that is where an apparent legal emptiness exists, or, at least, large differences among communities.

Broadly speaking, general regulation boundaries prohibit on-site betting to minors and those registered in Access Limitation (Conde, 2018), but beyond these, there are no other strict rules to follow by Autonomous Communities. It all depends on its management, which is substantially influenced by the ruling political force.

Bringing into focus on our work's frame, Barcelona held 47 establishments in the lasts of 2017, while Madrid gave shelter to 308. Comparing the dimensions and characteristics of the two largest cities of Spain, figures obtained should not make sense if it is not due to the differences in the regulatory landscape.

Accordingly, Madrid, Extremadura and Asturias are the only Autonomous Communities that have not developed a planning regarding this issue (Pascual and Grasso, 2018). Is for that reason that some streets of the Spanish capital are filled with more and more betting establishments with no possibility to take immediate action.

On the other hand, Cataluña is ruled on the basis of the Article 3 of 14/1984 Law of 20 March – which is currently developed through the 240/2004 Decree of 30 March – the approval and authorization of game and betting houses. Thereby, in accordance with this decree and involving our research, the planning of betting establishments in Cataluña is defined by the article 14 as follows: “The number of authorizations of betting houses (or B type) is limited to 126<sup>1</sup> for the entire territory of Cataluña” (Departament de la Vicepresidència i d’Economia i Hisenda, Generalitat de Catalunya, 2018).

### ***Justification of the project***

One of the main concerns about the Spanish gambling market is the proliferation of gambling premises across the country, with 425 new premises opened between 2016 and 2017 for a total of 2.896. Its increasing demand is mainly due to the offer of new technologically sophisticated machines and to sports bets between young man (Gómez, J. and Carlos, L., 2018), being *Sportium* and *Codere* the most important private operators.

---

<sup>1</sup> Note: Regarding betting premises, there are currently 127 in compliance with sentence 804/2015, of December 18. Tribunal Superior de Justícia de Catalunya.

As stated above, each Autonomous Community bears with the competencies regarding on-site betting activities and thus game halls are heterogeneously distributed across the country. Madrid's increase from 185 premises to 318 within 3 years has raised awareness and driven media attention to the topic as most of them have been established nearby school premises and in low-income neighborhoods.

In a press article from *El Confidencial*, the authors reveal that the growth of the number of gambling premises in Madrid between 2014 and 2017 occurred mainly in low income-level districts, where average income was below €25,000 per year (Pascual, A. and Grasso, D., 2018).

In this paper, we intend to go beyond a mere correlation and seek for causality. Our main focus remains on determining whether income is a determinant factor, while adjusted or not by other control variables, as it seems to be the case of Madrid. By doing this, we are trying to infer the rationale behind how big gambling corporations establish their premises with the ultimate objective of increasing their profits.

The decentralization of competences regarding betting regulation towards Autonomous Communities allows us to compare a highly regulated environment (*Cataluña*) against a more laissez-faire one (*Comunidad de Madrid*) that might enable to spot significant differences between both cities regarding our hypothesis. However, we could not be completely certain that regulation is the only cause if differences were appreciated.

Concluding, the main purpose of this study is to bear out whether gambling operators establish their physical premises where income is low within Barcelona and Madrid. In order to test this hypothesis, we use different regression models to corroborate or reject our initial suspicions rooted on *El Confidencial* article findings.

## LITERATURE REVIEW

As we have seen gambling industry fast-paced growth is expected to sustain and improve in the upcoming years. The improvements the sector is undertaking regarding its technologies and users' data analytics configure some of its most valuable assets as its overall figures do nothing but to increase. Such numbers have inspired curiosity from researchers into the spatial distribution of gambling and its expansion as well as the impact they have on society.

Several studies provide evidence that countries' tradition and advanced landscape regarding gambling are key to its development so that the spatial patterns of its distribution are significantly influenced by the social and cultural characteristics of its environment (Raento and Schwartz, 2011).

In addition, cities act as a cornerstone for the spatial distribution of the gambling landscape given that it tends to cluster in huge urban agglomerations. If the construction can be successfully integrated inside the neighborhood structure, in dimensions such as sustainability and connectivity, it can lead to the creation of tourism amenities amongst others (Klebanow and Gallaway, 2015).

Therefore, with reference to this study, the main characteristic evaluated is income as explanatory variable of gambling premises' location. Previous geospatial analysis of premises records in Britain found a significant relationship between gambling machine density and socio-economic deprivation. Those areas who had higher machine density per zone scored higher in income deprivation, higher in economically inactive people and a younger age profile than other areas (Wardle et al., 2013).

Income sustains a positive and linear relationship regarding gambling prevalence, the number of gambling activities practiced and attitudes towards this activity. However, as we will explain later the distribution of gambling premises does not follow this linear model. In fact, when assessing for problem and moderate risk gamblers according to Problem Gambling Severity Index (PGSI), the richest are amongst the least likely and conversely, the poorest are amongst the most likely, to punctuate in that way (Orford et al., 2010).

Another study found that social and demographic factors inducing problem gambling are associated, amongst others, with low income, low qualifications and residence in deprived neighborhoods (Abbott et al. 2015), where the lack of leisure opportunities for disadvantaged groups prompts that most of the gambling takes place on-site whereas in richer areas gambling is preferably done online (Fiedor et al., 2017). Therefore, this clearly affects the spatial distribution of gambling and influences inhabitants' behavior.

Research conducted in Nevada for the period 1900-1960 concluded that living within 10 miles or less of a gambling venue doubled the risk of problem gambling and likewise, individuals living in a disadvantaged neighborhood had a 90% odds' increase of becoming a pathological gambler (Satish and Kenneth, 2015).

Similarly, evidence of gambling proliferation in socio-economically deprived areas showed that the lower-middle classes with lower educational level gambled the most. Moreover, correlations to the gamblers' employment status, age and immigrant background would also support the idea that low socio-economically educated individuals gamble more (Beckert and Lutter, 2009).

Moreover, it is important to consider the spatial autocorrelation between the different areas studied, which enables to determine if there is interaction between neighborhoods due to their proximity. Income differences favor the clustering and settlement of poverty spaces which are also subject to social exclusion and insufficient services provision, that in consequence makes it difficult to turn around the situation (Aguilar and López, 2016). Thus, income not only would affect

the spatial distribution of gambling premises but also it would induce interplay between such low income areas.

However, an area gambling propagation may also largely depend on the objectives both public sector and the gambling industry have in different administration levels and therefore why different gambling policies coexist under the same country (Fijnaut and Littler, 2007).

In addition, regulation policies could be enhanced given the impact they have on society by complex statistical models to help regulators in decisions concerning licenses approval, in assessing the impact of measures such as gambling machines reductions on health and in identifying vulnerable hotspots (Markham, Doran and Young, 2014).

## METHODOLOGY

### *Generalized Linear Models (GLMs)*

Since linear models are not useful in cases where the responses variables are not normally distributed, Generalized Linear Models (GLMs) are the appropriate tool to predict the expected value of a given unknown quantity. (McCullag, P & Nelder, P., 1989; Agresti, A., 2007; Eberly College of Science, 2018; Rodríguez, G., 2007)

GLMs are an extension of classical linear models to fix their limitations. They are formed by three components:

- The random component, that assumes a probability distribution in order to identify the response variable Y.
- The systematic component, that specifies all explanatory variables with a linear combination of these variables called linear predictor ( $\alpha + \beta_1x_1 + \dots + \beta_kx_k$ ).
- The link function, that (with  $\mu$  being the expected value of Y) specifies a function  $g(\cdot)$  that relates  $\mu$  to the linear predictor as:

$$g(\mu) = \alpha + \beta_1x_1 + \dots + \beta_kx_k$$

where,  $g(\cdot)$  is the link and is used to the random and systematic components.

The link function is useful because a GLM does not assume a linear relationship between the dependent variable and the independent variables, but it does assume linear relationship between the transformed response in terms of the link function and the explanatory variables. Later we will see which link function is appropriated for our case analysis

In our study, we will need to use this model since our random component (i.e. number of betting houses per neighborhood) follows a Poisson distribution. Next, we will explain what is the Poisson distribution and what implications it has for the modeling of our regressions.

## ***Poisson distribution***

The Poisson distribution is used when count data behaves in this way (McCullag, P & Nelder, P., 1989; Agresti, A., 2007; Rodríguez, G., 2007):

- The probability of at least one observation of the event in a given time (or space) interval increases proportionally with the length of that interval.
- The probability of more than one observation in a very small time (or space) interval is extremely rare.
- The observations are mutually independent in disjoint time (or space) intervals.

The mean of the Poisson distribution is  $\mu = \lambda$ , being  $\lambda$  the number of observations of the event per unit of time (or space). The variance of this distribution is  $\sigma^2 = \lambda$ , so the Poisson distribution tells us that the variance increases as the mean does.

## **Link-log function**

Having a linear link function ( $g(\mu) = a + \beta_1 x_1 + \dots + \beta_k x_k$ ) for the case of the Poisson, we see that the right-hand of the equation can adopt any real value while the mean of the left-side cannot be negative (McCullag, P & Nelder, P., 1989; Agresti, A., 2007; Eberly College of Science, 2018) . A solution to this problem is using the logarithm of the mean:

$$\log(\mu_i) = x_i' \beta.$$

Now, the  $\beta$  coefficient tells us “the expected change in the log of the mean per unit change in the predictor  $x_i$ ” (Rodríguez, G., 2007). Using a link log function, we solve the problem of having negative values in the mean and, moreover, it fits better for the count data since their effects are, in general, multiplicative.

## **Offset variable**

Sometimes it is useful to form a rate for the count data (Agresti, A., 2007). It happens when the events that we are analyzing occur over, for example, space, so that then we can use an index to make models focus on the rate at which the events occur, i.e. dividing the count data by an index. Therefore, if we defined the index as  $t$ , the rate will be  $Y/t$ , or  $\mu/T$  (because  $\mu = E(Y)$ ). Then, a loglinear model for the expected rate has the following form:

$$\begin{aligned} \log(\mu/t) &= a + \beta_x, \text{ or} \\ \log \mu &= \log t + a + \beta_x \end{aligned}$$

In our regression models we will use the number of premises per neighborhood as an offset variable to the random component. Consequently, our count data will be comparable for every neighborhood, since the effect of having areas with different number of potential locations will be canceled.



## Akaike's Information Criteria (AIC)

Akaike's Information Criteria (AIC) is a measure of relative goodness of fit (Agresti, A., 2007; Rodríguez, G., 2007). It means that is a tool used to choose between a different set of models. The mathematical expression of the AIC is:

$$AIC = -2(\log \text{likelihood} - \text{number of parameters in model})$$

Where, the log likelihood in our case (a Poisson regression) is:

$$l(\mu) = -n\mu - \sum \ln(x_j!) + \ln(\mu)\sum x_j$$

The most parsimonious models are those with the lowest AIC. It takes into account two factors of the models: How close are the predicted values to the real ones and the number of parameters in the model, penalizing those with many parameters.

## Spatial autocorrelation

Spatial Autoregressive (SAR) models are used with data that has observations on spatial units or areas, in our case neighborhoods. They are an extension of the linear regression model that let outcomes be altered by outcomes, covariates and errors from adjoining areas.

Using the spatial jargon, SAR models can contain spatial lags, both of the outcome variable and the covariates, and spatially autoregressive errors. All these concepts come from time-series literature, where it is obvious that in many situations previous outcomes have an effect on the following ones.

An autoregressive process in time series can be expressed as the following:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \epsilon_t$$

Where,  $y_{t-1}$  is interpreted as the lag of  $y$ . And, using vector notation, the lag of  $y$  can be written as  $L.y$ .

$$y = \beta_0 + \beta_1 L.y + \epsilon$$

As we previously stated, autoregressive models can also have autoregressive errors:

$$u = \rho L.u + \epsilon$$

And thus, the autoregressive model becomes:

$$y = \beta_0 + \beta_1 L.y + (1 - \rho L.)^{-1} \epsilon$$

When the previous model gets translated to a SAR model, the lag of  $y$  becomes an  $N \times N$  matrix  $W$ , which is called the spatial weighting matrix. The corresponding SAR model with autoregressive errors becomes:

$$y = \beta_0 + \beta_1 W y + (I - \rho W)^{-1} \epsilon$$

Where  $(I - \rho W)^{-1} \epsilon$  is the spatial autoregressive error.

### Spatial weighting matrix (W)

The values in the Spatial Weighting Matrix W (StataCorp., 2017) represent the spatial relationship between areas. A W matrix can be used to allow both nearby outcomes and covariates to affect outcomes, and to take into account the effect autoregressive errors have on outcomes.

The most common way to construct a W matrix is through the usage of shapefiles, which, in our case, were obtained from the internet (Departament de Territori i Sostenibilitat., 2018) (Agencia para la Administración Digital de la Comunidad de Madrid, 2016).

In order to create our W matrices, one for Barcelona and one for Madrid, we have assumed that only adjacent neighborhoods affect each other. We have done it with the Stata command of: `spmatrices create contiguity W`. Previously, we had to mix the shapefiles of each city to its respective Stata dataset.

Moran's *I* test (Alexander, N., 2011) is the most commonly used technique to test spatial variability. It measures the spatial autocorrelation of the explanatory variable. It can be expressed as:

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Where *n* means the total number of spatial units, in our case the number of neighborhoods (73 in Barcelona and 128 in Madrid);  $w_{ij}$  is the weight between location *i* and *j* contained in the W matrix;  $y_i$ ;  $y_j$ , represent a concrete attribute, in our case, the number of betting houses, in locations *i* and *j* respectively. And  $\bar{y}$  in our case will be the average of betting houses per neighborhood.

The Moran's *I* statistic's range goes from -1 to +1. A value near to zero indicates spatially independence of the distribution. The null hypothesis is that the explanatory variable is spatially independent (i.e. Moran's *I* statistic being close to 0). A Z-score is normally used to indicate the significance of the test.

$$Z(I) = \frac{I - E(I)}{\sqrt{Var(I)}}$$

Where E(I) is the expectation of the statistic and Var(I) is its variance.

## DATA

### *Data resource*

#### *Response Variable*

The response variable used in the regression models, number of betting houses per neighborhood, corresponds to November 2017. Concretely, Barcelona's data was identified through the census of authorized betting establishments, administered by *Direcció General de Tributs i Joc, Generalitat de Catalunya*.

Out of 34 recognized establishments settled in Barcelona, another 13 betting houses managed by the market giants –Sportium and Codere- were found on their respective websites and included in the sample. Then, each location was checked and categorized both by district and neighborhood with the aim of obtaining a larger sample and, therefore, being able to carry out a more detailed and accurate analysis.

Equally important, Madrid's betting houses data was found in the municipal census of total establishments classified by activity. The monthly database regarding November 2017 contained a list of 161,097 establishments, from which those classified by 920002 epigraph – Chance games and private management bets – were chosen and categorized by district and neighborhood, as in Barcelona's case.

#### *Explanatory Variable*

Since the main objective of the statistical analysis is to find suggestive evidences relating betting houses' distribution with bettors' income, the selected explanatory variable was Household Average Income per district and neighborhood. In this case, and due to data access issues, the categorization of the variable is slightly different between Barcelona and Madrid samples. So, although results will be virtually unaffected, the interpretation of them will be somewhat different. In this way, both datasets will be previously described and clarified in order to avoid misunderstandings.

To depict Barcelona's market landscape, Household Disposable Income per Capita (HDIC) – amount of income available to resident households for consumption and savings, after tax deduction or fixed capital consumption – was picked in the form of an index, where Barcelona = 100. Hence, it shows the position of each neighborhood in relation to the city ordered from highest to lowest. The data was collected from the Department of Statistics of Barcelona, whose research dates back to 1<sup>st</sup> January 2016.

On the other hand, Madrid's available data regarding HDIC dates 1<sup>st</sup> January 2015 and was collected from an Urban Audit done by the Spanish National Institute of Statistics. Contrasting with Barcelona, figures were obtained in absolute numbers rather than indexed, hence our analysis should take this into account when interpreting.

## *Control Variables*

While looking for control variables to enhance the results' accuracy, it is required to select those ones that shows a relation both with  $x$  and  $y$  variable. Instead, they should not be correlated with each other. To this end, our statistical model includes two control variables explained below:

### *Education Level*

Remarkable literature about this field studied the relationship between bettors' social status and the propensity of being addicted to gambling. Indeed, some authors highlight the weight of education level on their model, since it is substantially correlated with household income. Theoretically, those with lower education tend to be poorer and, hence, more likely to become addicted (Hahmann and Matheson, 2017). Therefore, sports betting houses locate themselves nearby aiming at taking advantage of their addiction.

Our model simplifies the educational level in two categories: below or above standards. In the case of Barcelona, our model's high education standard was settled from *Bachillerato*, high school or equivalents. Accordingly, the percentage included in the dashboards comprises the percentage of working age population below standards; said, with obligatory education or less than that. Statistics were collected from the Department of Statistics of Barcelona for December 2016.

Secondly, education level classification for Madrid is slightly different, but our standard remains the same: low level of studies standard is set for those students with a lower degree than BUP, *Bachillerato Superior*, high school or equivalents. The dataset derives from Madrid's data bank and dates back to 1<sup>st</sup> January 2017.

### *Immigration ratio per 1000 inhabitants*

After visualizing the clustered distribution of betting houses across Barcelona and Madrid, we identified some factors that could incidence our model. One of them was immigration ratio per neighborhood, since some studies explain that immigrant neighborhoods are one of the main target of betting houses (Ordoñez, 2018). In addition, average immigration shows a negative correlation with income.

Both case studies date back the same period – Barcelona, December 2016, while Madrid, January 2017 –. In the first scenario, the number of foreign population by neighborhood was obtained via the Department of Statistics of Barcelona in absolute figures, and then computed – division over population of the same period – in order to obtain the ratio of immigrants per each 1000 inhabitants. Data from Madrid was also obtained from municipal population census and following the same process.

## Premises census – Offset variable

Finally, the model also includes neighborhood's number of premises as an offset variable. The aim is to normalize the distribution of sports betting houses by the number of potential locations they have in each area of study.

Both Barcelona's and Madrid's data dates back to 2017. Premises selected were those corresponding to ground floor and, in Barcelona's case, ignoring those dedicated to residential and parking uses.

## Data description

### Barcelona

Variable	Betting houses(*)	HDI Index	Immigration Ratio	Low Education Level	Premises
Mean	0.0002259	93.0137	50.17671	47.60411	2850.397
95% CI	[0.000160, 0.0003004]	[83.28663, 102.7408]	[45.70087, 54.65256]	[43.65492, 51.5533]	[2224.98, 3475.815]
Std. Dev.	1.01	41.69029	19.18352	16.92628	2680.545
Min-Max	0-5	[34.3,242.4]	[18.6,140.8]	[15.4,81.9]	[81,15588]
Median	0	84.4	45.7	47.6	2169
IQR	[0,1]	[63.7,105]	[39.9,56.9]	[34.7,60.2]	[1050,3859]

Table 1. Descriptive analysis of variables of interest in Barcelona.

-Betting houses: here the mean and the 95% CI are not expressed as a count variable, i.e. it is not a discrete variable, because the offset variable number of premises are taken into account. Therefore, we interpret the mean (0.0002259) as the rate of betting houses per local.

For the rest of the statistic measures, we have used the absolute value of this variable, resulting in a Standard Deviation of 1.01, a minimum number of betting houses of 0 (that can be found in 44 neighborhoods) and a maximum of 5.

In Fig.1., we see that the variable has an exponential decay, with only three areas with more than two betting houses.

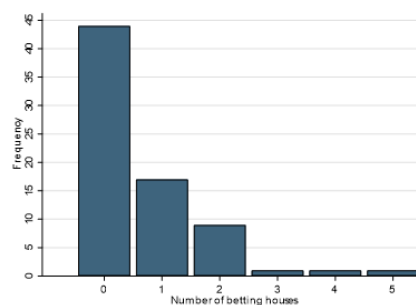


Fig.1. Histogram of the number of betting houses per neighborhood in Barcelona.

It can be observed in Fig.2. that betting houses cluster in the districts of *Ciutat Vella* and *l'Eixample*, with an important quantity also located in *Sants - Montjuïc*.

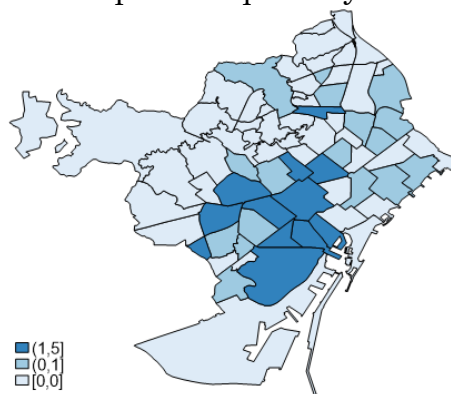


Fig.2. Number of betting houses by neighborhood in Barcelona.

-Household Disposable Income (HDI) Index: since we are dealing with data from neighborhoods and not from population, our mean is not 100. Therefore, we can extrapolate that neighborhoods with more population are the wealthiest ones.

The Standard Deviation is very high and, looking at the median and the IQR, we notice that the majority of the neighborhoods have a HDI Index lower than 100.

In Fig.3. we observe that the wealthiest neighborhoods are the ones from the district of *Sarrià – Sant Gervasi*, *l'Eixample* and *Les Corts*, whilst some other high HDI Index neighborhoods are dispersed over the city.

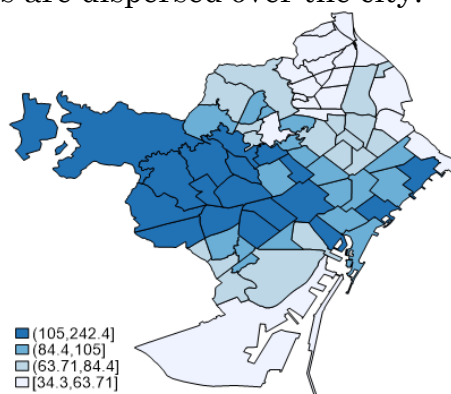


Fig.3. Household Disposable Income Index by neighborhood in Barcelona.

-Immigration ratio: expressed as the number of immigrants per 1,000 inhabitants, its mean is 50.18, although is a very dispersed variable with a standard deviation of 19.18.

The highest levels of immigration can be found in the districts of *Ciutat Vella*, *l'Eixample*, some points of *Sarrià – Sant Gervasi*, *Sants – Montjuïc* and *Sant Martí*. Lower levels of immigration can be found in the neighborhood of *La Marina del Prat Vermell - Zona Franca* and the district of *Nou Barris*.

-Low Educational Level: which is the index of people with No Education or only the Mandatory Level over the total population, it has a mean of 47.6% and it is relatively disperse.

The lowest educated neighborhood is *La Marina del Prat Vermell - Zona Franca* (81.9%), followed by some from the districts of *Nou Barris* and *Sant Martí*. The highest educated neighborhoods are, in general, located in the *Sarrià – Sant Gervasi* district.

-Premises: This variable is our offset variable. We observe a wide range in the minimum and maximum values. However, we see that there are strange values since the CI of the mean and the IQR do not disclose such differences.

The touristic and more populated areas are the ones with more premises, e.g. the districts of *Ciutat Vella*, *l'Eixample*, *Les Corts* and *Sarrià – Sant Gervasi*. The districts of *Nou Barris* and *Horta – Guinardó* have the lowest values of premises in their neighborhoods (Fig.4.).

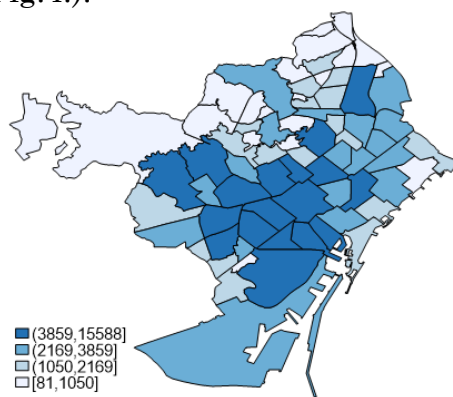


Fig.4. Number of premises by neighborhood in Barcelona.

### Madrid

Variable	Betting houses(*)	HDI Index	Immigration Ratio	Low Education Level	Premises
<b>Mean</b>	0.0031515	42585.41	120.56	45.61	763.52
<b>95% CI</b>	[0.0028093, 0.0035239]	[39167.34, 46003.49]	[110.92, 130.21]	[42.18849, 49.02167]	[668.2614, 858.7855]
<b>Std. Dev.</b>	2.421809	19542.51	55.14	19.53	544.65
<b>Min-Max</b>	[0,11]	[19587.09, 112320.8]	[31.404, 297.784]	[16.73, 82.68]	[15, 2697]
<b>Median</b>	2	35932.7	108.559	43.61	682.5
<b>IQR</b>	[0,4]	[28386.6, 48385]	[81.2025, 150.97]	[26.9, 61.89]	[318, 1056.5]

Table 2. Descriptive analysis of variables of interest in Madrid.

-Betting houses: the processing of this data for the calculation of the mean and its CI is the same as for Barcelona, giving us 0.0031515 and [0.0028093, 0.0035239] respectively.

Fig.5 shows a decreasing shape with lots of neighborhoods with 0 to 4 betting houses (IQR: [0, 4]) and unusual cases of neighborhoods with 5 or more betting houses.

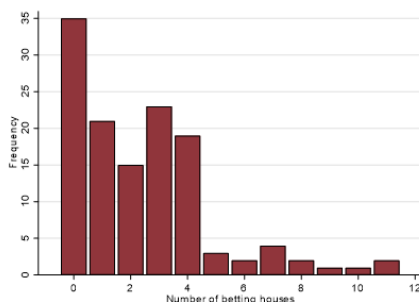


Fig.5. Histogram of the number of betting houses per neighborhood in Madrid.

Most of the neighborhoods have more than one betting house and less than four. Some neighborhoods in *Fuencarral – El Prado*, *Moncloa –*, *Aravaca*, *Hortaleza* and *Barajas* don't have betting houses, meanwhile the districts of *Vicalvaro*, *La Latina*, *Tetuán* and *Puente Vallecas* are the ones with a higher concentration (Fig.6).

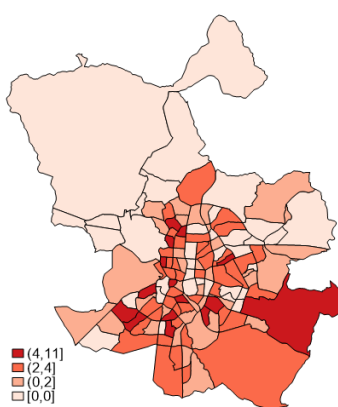


Fig.6. Number of betting houses by neighborhood in Madrid.

-HDI: here we are dealing with absolute values. The variable is very dispersed since the Standard Deviation has a high value. The median shows us that there are more neighborhoods with lower income than the mean.

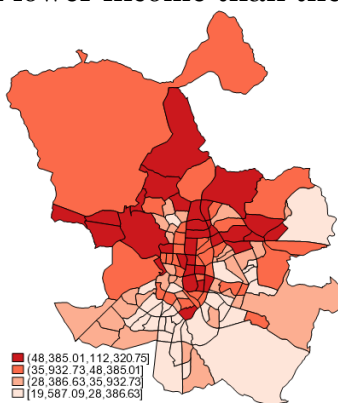


Fig.7. Household Disposable Income Index by neighborhood in Madrid.



In Fig.7. we see that the wealthier areas are in the districts of *Salamanca*, *Chamberí*, *Moncloa – Aravaca*, *Fuencarral – El Prado* and *Chamartín*. The lower income ones are in *Puente Vallecas*, *la Villa de Vallecas*, *Villa Verde*, *Usera* and *Carabanchel*.

-The Immigration ratio: here we have the number of immigrant population over 10,000 inhabitants. Here we see that there are more neighborhoods with a lower index than the mean (median: 108.6, mean: 120.56).

The neighborhoods with more immigration are located in the districts of *Villa de Vallecas*, *Villaverde*, *Usera*, *Carabanchel* and some neighborhoods (e.g. *Sol*: 246) of the centre of the city with huge differences between them.

-Low Education Level: this variable has a mean of 45.61% and is quite dispersed. The districts of *Salamanca* and *Chamartin* record the lowest levels for this variable and the neighborhoods of the districts of *Puente Vallecas* and *Usera* score the highest levels.

-Premises: the number of premises has a mean of 763.52 but its median is lower (682.5). The first impression is that the number of little neighborhoods with fewer premises is high, but in fact the neighborhoods with fewer premises are located in *Fuencarral – El Prado*, *Moncloa – Aravaca* and *Barajas*, districts whose neighborhoods are big in extension. Thus, there are specific areas (*Casco Histórico de Vallecas*, *Villaverde Alto*, *Universidad* and *Embajadores*) that considerably increase the mean.

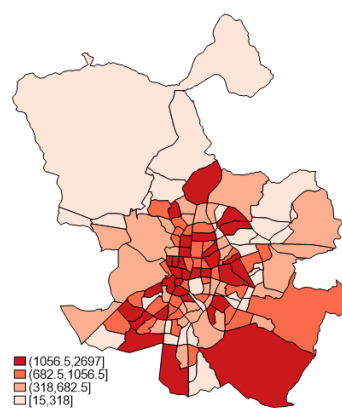


Fig.8. Number of premises by neighborhood in Madrid.

## RESULTS & DISCUSSION

### *Results*

#### *Regressive models*

In order to identify the impact that household disposable income may have on the number of betting houses in a neighborhood, and whether or not this impact is significant, we have created four different regression models, both for Madrid and Barcelona.

Each model satisfies the conditions specified in the methodology part. They all have in common that are Generalized Linear Models (GLMs) of the Poisson family and the usage of the number of non-residential premises per neighborhood as an offset variable.

Model 1 uses no control variable. Model 2 uses the immigration ratio as a control variable. In Model 3, low education level is used as the control variable. Finally, in Model 4 both immigration ratio and low education level are used as control variables.

The results for the estimated coefficients, the robust standard errors, the 95% confidence intervals, the p-values and the AIC are the following:

<b>BARCELONA</b>	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>	<b>Model 4</b>
<b>Coefficient</b>	-.0014014	-.0016218	-.006132	-.0058152
<b>Robust Std. Err.</b>	.0025761	.0025146	.0056779	.0057008
<b>95% CI</b>	[-.0064504, .0036476]	[-.0065503, .0033067]	[-.0172604, .0049964]	[-.0169886, .0053581]
<b>P&gt; z </b>	0.586	0.519	0.280	0.308
<b>AIC</b>	1.673969	1.687097	1.694666	1.709419

*Table 3. Regression models' results in Barcelona.*

In Barcelona we can identify that in all four models the estimated coefficient takes negative values. However, no one of the p-values is significant neither at 99%, 95%, nor 90% of confidence. Consequently, the range of the 95% confidence interval of the coefficient includes zero in each model. Additionally, Model 1 is the model which has the lowest Akaike's Information Criteria value (AIC), being the most parsimonious.

It is also worth mentioning that in Barcelona the immigration ratio happens to have a statistically significant effect on betting houses per neighborhood in both Model 2, with 95% of confidence, and in Model 4, with 90% of confidence. However, this is not the scope of our project.

<b>MADRID</b>	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>	<b>Model 4</b>
<b>Coefficient</b>	-.0000275	-.0000249	-.0000284	-.0000247
<b>Robust Std. Err.</b>	5.28e-06	6.31e-06	.0000114	.0000107
<b>95% CI</b>	[-.0000379, -.0000172]	[-.0000373, -.0000125]	[-.0000507, -6.14e-06]	[-.0000457, -3.78e-06]
<b>P&gt; z </b>	P<0.01	P<0.01	0.012	0.021
<b>AIC</b>	3.372439	3.382904	3.387909	3.398524

*Table 4. Regression models' results in Madrid.*

In Madrid we identify again that in all four models the estimated coefficient takes negative values. In this case being statistically significant at 99% of confidence in both Model 1 and Model 2; and at 95% of confidence in Model 3 and Model 4. Therefore, the range of the 95% confidence interval of the coefficient comprises only values below zero in every model. Moreover, Model 1 is again the model which has the lowest AIC value, thus being the most parsimonious.

In Madrid no one of the two control variables proof to have a significant impact on the number of betting houses in any of the models.

#### *Spatial autocorrelation*

The Moran's *I* Test has been taken for both cities Madrid and Barcelona. In each case it has been tested whether the errors from the simple linear regression between the number of betting houses and the respective measure of household income are independent and identically distributed (i.i.d.) or not, from one neighborhood to its contingent areas.

In Barcelona the Moran's *I* test rejects the null hypothesis of the errors being i.i.d with a p-value lower than 0.01. Contrarily, in the city of Madrid the p-value is 35.5%, thus not rejecting the null hypothesis at any one of the most used significance level.

#### ***Discussion***

The results have shown that in both cities it is projected to be present an inverse relationship between the number of betting houses per neighborhood and the economic status of its residents. Yet, it is only in the city of Madrid where this relationship has high levels of statistical significance.

The interpretation of the coefficients should be understood as the average change of the logarithm of the number of betting houses per neighborhood, offsetted by the number of non-residential premises, due to a one-unit increase of the measure of household income. As Model 1 proves to be the most parsimonious in each city we should use its coefficients to make this inference.

On the other hand, Moran's I test indicates that in Barcelona the areas of study are not independent and identically distributed. However, in the case of Madrid this affirmation cannot be done at any relevant significance level.

Linking the results and interpretation of our analysis with what we could have expected from pre-existing literature, we can conclude that Madrid is a perfect example regarding the spatial distribution of gambling premises in low-income neighborhoods. However, neither education nor immigration coefficients are significant so to label such areas as socio-economically deprived, where low-income, under-educated and also higher immigrants population percentage inhabitants would coexist.

Interestingly, in Barcelona, where strong regulation ruling the Autonomous Community limits gambling corporations' development, the significance of the results do not allow for the same conclusion.

Thus, we sustain the suspicion that gambling policies differences in important factors, such as licenses limitation and location prohibitions, account for the spatial distribution difference between both cities. Limitations on the number of betting houses authorizations might be causing tradeoffs that are somehow diluting the incentives to locate in low-income areas.

It is also worth mentioning that in the city of Barcelona immigration impact has been found to be statistically significant, allowing for future studies that are further of this one.

Moreover, it does not seem that spatial autocorrelation takes place in Madrid. The errors from the simple linear regression model used by Stata to perform the Moran's I test flaw the results and thus their interpretation because our random variable follows a Poisson distribution. On the other hand, in Barcelona this autocorrelation is more likely to take place although this result should not be taken as a consistent one.

Consequently, we cannot confidently proof that there is interaction between neighborhoods in both cities, and especially between low-income ones, as literature suggests a tendency to cluster between poverty spaces.

## CONCLUSION

Gambling premises are encountered, on average, to be located in low-income areas both in Barcelona and Madrid. Yet, this relationship only proofs to be sound for the latter. Besides, spatial autocorrelation driven analysis has only been found to be significant in Barcelona. All in all, everything suggests that a plausible and probable source of such results' disparities could be the gambling regulation enforced by Autonomous Communities' government.

Implications of the results revealed in this article should be subject of study by *Comunidad de Madrid*, *Extremadura* and *Asturias*' competent authorities. A

decision to take a stand and apply regulations regarding capacity limitations, taxation heterogeneity and location prohibitions, would clearly be accounted as having an important social impact and a key role in explaining spatial spread over time.

## REFERENCES

Abbott, M., Binde, P., Hodgins, D., Korn, D., Pereira, A., Volberg, R. & Williams, R. (2015): Conceptual Framework of Harmful Gambling: An International Collaboration. *Guelph: Ontario Problem Gambling Research Centre (OPGRC)*.

Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Aguilar, A. & López, F. (2016). Espacios de pobreza en la periferia urbana y suburbios interiores de la Ciudad de México. Las desventajas acumuladas. *Eure*, 42(125): 5-29.

Alexander, N. (2011). Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology. *Journal of the Royal Statistical Society*, 174 (2): 512.

Beckert, J. & Lutter, M. (2009). The inequality of fair play: Lottery gambling and social stratification in Germany. *European Sociological Review*, 25(4): 475–488.

Boletín Oficial del Estado. (2011). *BOE núm. 127, de 28/05/2011*. Retrieved November 26, from <https://bit.ly/2nK5vn8>

Conde, J. (2018). *Jóvenes ludópatas enganchados a casas de apuestas, una “epidemia” social difícil de controlar*. Extremadura, Spain: El Diario Extremadura. Retrieved November 26, from <https://bit.ly/2rnGRwc>

Departament d’Estadística i Difusió de Dades, Ajuntament de Barcelona. (2016). *Level of studies per district and neighborhood*. Retrieved November 26, from <https://bit.ly/2QfvFjQ>

Departament d’Estadística i Difusió de Dades, Ajuntament de Barcelona. (2016). *Indexed household disposable income per capita*. Retrieved November 26, from <https://bit.ly/2UdZG1W>

Departament d’Estadística i Difusió de Dades, Ajuntament de Barcelona. (2017). *Immigrants by Spanish or foreign nationality*. Retrieved November 26, from <https://bit.ly/1Of7eeT>

Departament de la Vicepresidència i d’Economia i Hisenda, Generalitat de Catalunya. (2018). *Article 3 of Law 15/1984 of March 20. Decree 240/2004 of March 30*. Retrieved November 26, from <https://bit.ly/2zLS9yN>

Dirección General de Ordenación del Juego. (2018). *URL Directory of operators*. Retrieved November 26, from <https://bit.ly/2zFTOWE>

Eberly College of Science of the Pennsylvania State University. (2018). *STAT 504, Analysis of Discrete Data*. Pennsylvania, US.: Eberly College of Science. Retrieved November 23 from: <https://bit.ly/2QfHsi4>

Fenez, M. (2016). *Global Gaming Outlook – the casino and online gaming market to 2015*. New York, UE.: PwC. Retrieved November 26, from <https://pwc.to/2r8zIzh>

Fernández, M. (2018). *El juego cambia de ciclo en España*. Madrid, España: Diario El País. Retrieved November 26, from <https://bit.ly/2uBEBm5>

Fiedor, D., Szcyrba, Z., Šery, M., Smolová, I. & Tousek, V. (2017): The spatial distribution of gambling and its economic benefits to municipalities in the Czech Republic. *Moravian Geographical Reports*, 25(2): 104-117.

Fijnaut, C. & Littler, A. (2007). *The Regulation of Gambling: European and National Perspectives*. Leiden: Martinus Nijhoff publishers.

Gómez, J. & Carlos, L. (2018). *Anuario del juego en España 2018 – Fundación Codere*. pp. 241-242.

Gómez, J., Cases, J., Gusano, G., Carlos, L. (2016). *Anuario del juego en España 2014-2015 – Fundación Codere*.

González-Espejo, P., López, D. (2018). *La Nueva Ley del Juego*. Madrid, España: Abogacía Uría Menéndez. Retrieved November 26, from <https://bit.ly/2zJ2taF>

Hahmann, T., Matheson, F. (2017). *Problem gambling and poverty*. *Gambling Research Exchange Ontario, Toronto*.

Humphreys, B., Pérez, L. (2012). Who bets on sports? Characteristics of Sports Bettors and the Consequences of Expanding Sports Betting Opportunities. *Estudios de Economía Aplicada*, 30(2): 579-598.

Jones, B. (2018). *Harmless Harm*. Columbus, Ohio, US.: Maryhaven, a Central Ohio's behavioral health & addiction treatment center. Retrieved November 26, from <https://bit.ly/2rmEQQF>

Klebanow, A. & Gallaway, S. (2015). *Casinos and the City. A White Paper on the History of Casino Development in Cities, Past and Current Trends, and Recommendations for Future Development*. Las Vegas: Global Market Advisors.

Markham, F., Doran, B. & Young, M. (2014). Estimating gambling venue catchments for impact assessment using a calibrated gravity model. *International Journal of Geographical Information Science*, 28(2): 326–342.

McCullag, P. & Nelder, P. (1989). *Generalized Linear Models, Second Edition*. Boca Raton: Chapman and Hall/CRC

Ordoñez, E. (2018). *Apuestas deportivas: engançados a un juego al que es imposible ganar*. Madrid, España: Reportajes CTXT. Retrieved November 26, from <https://bit.ly/2Rvm6tM>

Orford, J., Wardle, H., Griffiths, M., Sproston & K., Erens, B. (2010). The role of social factors in gambling: Evidence from the 2007 British Gambling Prevalence Survey. *Community, Work & Family*, 13(3): 257–271.

Pascual, A., Grasso, D., (2018). *Las salas de apuestas exprimen los barrios pobres de Madrid: crecen un 140% desde 2014*. Madrid, Spain: Revista El Confidencial. Retrieved November 26, from <https://bit.ly/2OLD0IX>

Portal de Datos abiertos del Ayuntamiento de Madrid (2017). *Epigraphs of economic activity*. Retrieved November 26, from <https://bit.ly/2QzYgQt>

Portal de Datos abiertos del Ayuntamiento de Madrid. (2015). *Average net income of households, Urban Audit*. Retrieved November 26, from <https://bit.ly/2SoW2As>

Portal de Datos abiertos del Ayuntamiento de Madrid. (2017). *Census of premises and their activities*. Retrieved November 26, from <https://bit.ly/2rmcIgz>

Portal de Datos abiertos del Ayuntamiento de Madrid. (2017). *Population aged 25 and over classified by level of studies per district and neighborhood*. Retrieved November 26, from <https://bit.ly/2p3Sf0b>

Portal de Datos abiertos del Ayuntamiento de Madrid. (2017). *Population classified by nationality (Spanish or foreign) and gender, by district and neighborhood*. Retrieved November 26, from <https://bit.ly/2p3Sf0b>

Raento, P. & Schwartz, D. (2011). *Gambling, Space, and Time: Shifting Boundaries and Cultures*. Las Vegas: University of Nevada Press.

Research and Markets (2017). *Gambling Market – Global Outlook and Forecast 2018-2023*. Dublin, Ireland: Research and Markets. Retrieved November 26, from <https://bit.ly/2E3AtBR>

Rodríguez, G. (2007). *Lecture Notes on Generalized Linear Models*. Princeton University.

Satish K. & Kenneth E. (2012). *RIA Reaching Others with Gambling Research*. Retrieved November 24, from <https://bit.ly/2UdjiTU>

StataCorp. (2017). *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC

Toca, G. (2017). *El juego en España: una apuesta de 35.000 millones de euros*. Madrid, Spain: Revista Forbes España. Retrieved November 26, from <https://bit.ly/2i4HMQG>

Wardle, Heather, Keily, Ruth, Astbury, Gaynor, and Reith, Gerda (2014) 'Risky places?': mapping gambling machine density and socio-economic deprivation. *Journal of Gambling Studies*, 30 (1). pp. 201-212.