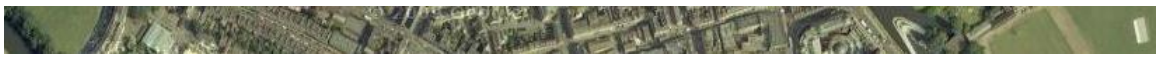
Geofutures

Gambling & Place Research Hub

Secondary Analysis of Machines Data

Examining the effect of proximity and concentration of B2 machines to gambling play



Prepared by Geofutures for

The Responsible Gambling Trust

Commercial and in confidence

Gaynor Astbury & Heather Wardle

29 March 2016

Geofutures Ltd
39 Powlett Road
Bath BA2 6QL
Tel: 01225 222382
www.geofutures.com

Contents

Executive summary

Aims and objectives

- This study aims to examine the relationship between spatial concentrations of B2 machines (commonly known as Fixed Odds Betting Terminals) in Licensed Betting Offices (LBOs) and gambling behaviour. It is an extension of previous research using survey data of loyalty card holders for LBOs and industry machine play data.
- LBOs with B2 machines cluster spatially in Great Britain and we have identified areas with greater spatial concentrations of LBOs with B2 machines.
- Having identified areas with greater concentrations of LBOs with B2 machines, we examined the relationship between living near differing **numbers of LBOs** or **a concentration of LBOs** and gambling behaviour and machine play.

Analysis and methods

- Using measures of **problem gambling** collected through a survey of people who held an LBO loyalty card, we examined if problem gambling rates differed according to whether or not someone lived in an area with a higher number or concentration of LBOs with B2 machines. We also looked at whether the **number of machine gambling sessions** and **the number of days** a person gambled on a machine varied by the number or spatial concentrations of LBOs with B2 machines.
- We used two methods to analyse spatial patterns of LBOs. The first was a simple count of the **number of LBOs** within 400m of a loyalty card holder's home. Our second method identified areas with **LBO concentrations**: 3 or more LBOs within 200m or 400m of one another.
- Analysis is based on people who held a loyalty card for one of three bookmakers in 2014. Loyalty card holders are **highly-engaged players**, therefore results cannot be generalised to all machine gamblers, but rather reflect patterns among highly engaged loyalty card customers.

Results

- There were no statistically significant differences in problem gambling prevalence, PGSI scores, the number of machine sessions or the number of days on which machines were played according to the number of LBOs someone had in their local area. However, a trend could be detected when combining rates of low risk, moderate risk and problem gambling (a PGSI score of 1 or more), which showed higher gambling prevalence rates in the combined risk groups compared to non-problem gamblers, when there are more LBOs in the local area.
- **Problem gambling and moderate risk prevalence rates were higher among those who lived in LBO concentration areas**. Machine players living in high concentration LBO areas tended to play slightly fewer machine sessions and play machines on a fewer number of days than those who did not live in high concentration areas, although the difference was small.
- Our results suggest that the **spatial configuration** of LBOs with B2 machines is important. Higher-density concentrations are associated with stronger patterns in gambling differences, than simple counts near to players' homes.

- This research provides the first evidence that problem gambling rates among machine players vary according to whether someone lives in proximity to a concentration of LBOs. However, these patterns and statistics can **indicate correlation, but they cannot determine causation.** Further research may look at the drivers behind these results in more detail.

1. Project background

In 2014 a consortium of NatCen Social Research, Featurespace, Geofutures and RTI International conducted a programme of research for the Responsible Gambling Trust (RGT) into Category B2 gaming machines in bookmakers. The research recognised the growing concern about B2 gaming machines (commonly known as Fixed Odds Betting Terminals or FOBTs) in Licensed Betting Offices (LBOs) in Great Britain. The main objective was to examine the extent to which industry data could be used to distinguish between harmful and non-harmful gaming machine play, as well as ascertaining what measures might limit harmful play without impacting on those who do not exhibit harmful behaviours.[1] We recommend reviewing the results of this project in conjunction with those reports.

A number of useful datasets were created for these previous research projects and in 2015 a range of secondary analysis projects were commissioned to explore this data further. This project is one of the additional projects commissioned.

1.1 Aims and objectives

This study examines whether there are any correlations between gambling behaviour and B2 machine gambling and the number and spatial concentration of LBOs with B2 machines in Great Britain (GB). The study uses the historic data gathered and used for the RGT machines research programme in 2014, and sits amongst several other research questions now being addressed by the wider consortium.

It aims to answer the questions:

a. **How do patterns of gambling relate to the number of B2 machine venues near a player's residence?**

b. **How do patterns of gambling relate to living near a concentration of B2 machine venues?**

1.2 Report structure

In Section 2 we provide details on the datasets used. In Section 3, we outline our approach and methodology. In Section 4 we present the results of this analysis and discuss the assumptions and known error margins in the analysis.

[1] The full body of research can be found at http://www.responsiblegamblingtrust.org.uk/commissioning/research/research-publications/

2. Datasets

Previous research created several datasets about gambling characteristics on B2 machines. These data consisted of a mixture of information recorded by B2 machines themselves and information collected through a survey of people who held a loyalty card for one of three bookmakers in 2014. Details of the data we have re-used for this analysis are outlined below.

2.1 The loyalty card survey

A survey of people who held a loyalty card for either Ladbrokes, William Hill or Paddy Power was carried out in 2014 to measure problem gambling prevalence among this group. A random probability sample was drawn from all loyalty card customers for Ladbrokes, William Hill and Paddy Power who had used their loyalty cards when playing B2 machines between September and November 2013, and had contact information available for phone interviews or email surveys. At the end of the survey, permission was asked to link survey responses with a players' loyalty card and machine data.

Overall, there were 180,542 loyalty cards of which 131,275 had some form of contact detail available, from which a random probability sample (n=47,268) was drawn. The sample was stratified, with those cards which had been used most often being oversampled to boost the number of gamblers who might be experiencing problems. All selected participants (n=47,268) were contacted by operators to inform them that they had been selected to participate and that NatCen Social Research would be contacting them unless they did not wish NatCen to do so. Overall, 902 people opted out of participating. This process also identified that 18,801 records had invalid contact details. The final issued sample size was 27,565.

Survey fieldwork was carried out between May and August 2014 via email and telephone. The questionnaire covered:

- engagement in a range of gambling activities in the past four weeks;
- frequency of gambling participation for each activity;
- use of loyalty cards;
- problem screening questions;
- attitudes to machines in bookmakers;
- motivations for playing machines in bookmakers;
- demographics;
- machine data linkage permission.

Overall, 4,727 loyalty card holders took part in the survey, with 4,001 (85%) agreeing that their survey responses could be linked to their loyalty card data which tracked their patterns of play on B2 machines in bookmakers. The represents a response rate of between 17%-19% (based on the random probability sample of 27,565 issued cases).

Address information was also requested within the survey questionnaire, from which players could be geo-referenced to the full unit postcode (the centre of around 15 adjacent addresses). When analysed, a total of 3,442 records were available with valid postcodes from within Great Britain for spatial analysis. These were derived from the following operators:

1,833 from Ladbrokes,
1,335 from William Hill,

274 from Paddy Power.

Survey data were weighted to correct both for non-response biases and to adjust for oversampling those who were more frequent machines players. However, there may still be some systematic biases in the data. For example, those who had valid contact details may be systematically different to those who did not, though the non-response weights accounted for this to a limited extent (see Wardle et al, 2014 for further details).

By their nature, loyalty card holders are those heavily engaged in gambling, and findings from this survey should not be extrapolated to all machine players. Indeed loyalty card survey participants have high rates of problem gambling and at-risk gambling compared to other national surveys, (Wardle et al, 2014b). With this in mind we can say that the study is representative of loyalty card holders who tend to be highly engaged in gambling.

Measuring problem gambling

The aim of the loyalty card survey was, mainly, to measure problem gambling prevalence among this group. This was measured by a series of nine survey questions called the Problem Gambling Severity Index (PGSI). These PGSI items include:

- betting more than a player can afford to lose;
- a need to gamble with increasing amounts of money;
- chasing losses;
- borrowing money or selling items to get money to gamble;
- feeling they had a problem with gambling;
- gambling causing health problems including stress and anxiety;
- people criticising gambling behavior;
- gambling causing financial problems for the player or the household;
- feeling guilty about the way that they gamble or what happens when they gamble.

Responses to each question ranged on a four-point scale from 'always' to 'never', which combined to produce a PGSI score with a maximum score of 27, (Wardle et al, 2014b). The PGSI then groups people into the following categories:

Table 1: PGSI aggregated groups.

PGSI classification category	**PGSI score**
Non-problem gambler / those who gamble without any difficulties	0
Low risk gambler	1-2
Moderate risk gambler	3-7
Problem gambler	8 or more

Overall, the loyalty card survey estimated that 23% of loyalty card holders were problem gamblers. This was an ***estimate*** of problem gambling rates among loyalty card survey participants which should not be extrapolated to all machine players due to the limitations outlined.

Around 45% of loyalty card survey participants said that playing machines in LBOs was their most frequent form of gambling activity, and all had gambled on machines in the last year.

However the PGSI score is measured based on all gambling activities, not just B2 machine play. Therefore, loyalty card survey participants who were identified as problem gamblers may experience problems with other forms of gambling as well as their B2 play, though it is notable that among problem gamblers in this survey 53% stated that machines were their most frequent form of activity.

In the sections that follow, we refer to this data as 'LCS participants', meaning data which is derived from the loyalty card survey (LCS) participants.

2.2 Operator data

In addition to the LCS participant data, a limited number of metrics were also gathered directly from operators relating to engagement in machine play. This was derived from the machine data held by operators, which was analysed and linked to loyalty cards active between September 2013 and June 2014. This data is useful because we had both the postcode of the loyalty card holders' home address and information about their machine play which was recorded against their loyalty card. What we do not have for these people were their problem gambling scores (as they were not survey participants).

After geocoding, this sample included;

- **Number of sessions** – 179,043 records from Ladbrokes, William Hill and Paddy Power, linked to loyalty cards.
- **Number of different days played on** – from the same sample, as above.

In the sections that follow, we refer to this operator data as 'machine play data', meaning this is the data we have derived direct from the operator's records.

2.3 B2 machine locations

B2 machine locations were provided by Inspired Gaming and Scientific Gaming, who supply all B2 machines to LBOs in Great Britain**.** These data included opening and closing dates, from which we could identify LBOs open in October 2014 (the date used for the preceding research). There is often more than one B2 machine in any one LBO and the research measures individual venues with at least one B2 machine rather than the count of B2 machines, as this was not known. Our study omits LBOs with no B2 machines present.

Since the preceding research, 82 records of LBO locations were deleted from this dataset as they were identified as potential duplicates. We have identified 9,272 LBOs open in October 2014 for this research.

LBO locations were georeferenced to the full unit postcode (the centre of around 15 adjacent addresses). There is often more than one venue within a postcode, and the analysis accounts for this.

2.4 Contextual datasets

Datasets were mapped using postcode information and geo-located using the Office for National Statistics Postcode Directory (ONSPD) extract at November 2015. This dataset gives a location for every full unit postcode, which represent around 15 adjacent addresses.

Further analysis to identify the surrounding population has been modelled using Census 2011 Output Areas, which are the smallest areas used to measure and map UK Census data.

Base mapping has also been used from the Ordnance Survey open datasets for geographic context.

All datasets are available as national open data.

3. Approach and methodology

3.1 Study context

This study focuses on the location of B2 machines, which mostly occur in LBOs in Great Britain. There are a small number of LBOs which do not have B2 machines, and so the approach represents patterns of LBOs with B2 machines rather than LBOs specifically. LBOs may have up to four machines, however the analysis does not account for the total number of machines as this was not known in the data provided to us. In this report, we call machines in bookmakers B2 machines for parsimony but recognise that these machine offer a range of games, including B3 and category C content as well.[2] However, B2 games (such as roulette) are the most popular games played on these machines.

Our study area is Great Britain since this was the geographic extent of the original research and thus data are available for Great Britain only. Within Great Britain we were attempting to model local-level patterns, and we assume this 'local' level to capture the walking distance from player residences to services, rather than city and regional wide trends. We also assume that in measuring the 'concentration of LBOs with B2 machines' we are trying to capture several LBOs with B2 machines on adjacent streets or a similar level, rather than capturing a concentration that may encapsulate an entire town or city. Our choice of modeling parameters at the local level reflect this scale of analysis.

One caveat is that the available data cannot identify the machines or LBOs at which players are gambling; individuals may have a large number of machines in their local area but may not necessarily be playing at these venues. However, previous research showed that those who gambled most frequently on machines tended, on average, to travel shorter distances to their most frequent LBO. This suggests that very frequent machine players are more likely to be visiting LBOs locally to their home residence (Astbury & Thurstain-Goodwin, 2015). A further caveat is that we are locating players by their residences, and the research cannot account for proximity to LBOs from players at their work or other regularly-visited places away from their home.

There are several ways to investigate the relationship between gambling behaviour and the location of LBOs with B2 machines. In this report, we use two approaches in order to compare the results and understand any trends that may be occurring. These approaches are discussed in the sections that follow.

3.2 Assessing the sample

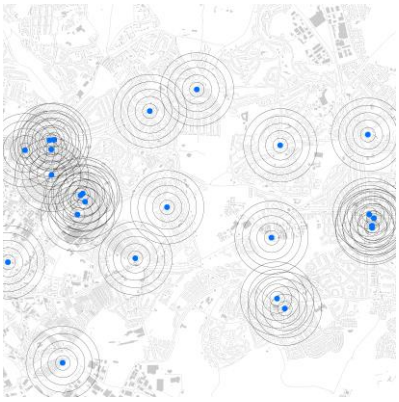
Our spatial analysis assumes several geographic patterns in the data. Firstly we presume that there is a variation in how LBOs are geographically spread across the country, with concentrations or 'clusters' of LBOs occurring in some areas and not others, which may affect gambling behaviours. We have used spatial statistics to ascertain whether LBOs exhibit these patterns.

[2] In Great Britain, gaming machines are split into different types of categories based on the stake and prize limits offered. B2 games have a maximum stake of £100 and prize of £500 whereas B3 games have a maximum stake of £2 and prize of £500.

Firstly we have used the Ripley's K statistic to compare the geographic patterns of LBOs and players compared to all residential patterns in Great Britain, to identify clusters of LBOs and players within the overall population.

The Ripley's K statistic measures the degree of clustering of LBOs by calculating the 'intensity' of points across a study area, using circles of progressively larger size. For a detailed discussion of this statistic, please see Appendix A. Results are discussed in section 4.1.

Figure 1: illustrative Ripley's K calculation.



Secondly, the LCS participants and machine play data may exhibit some bias in where people are under and over-represented in the country. We are therefore interested in the geographic spread of players across the country to assess the geographic representativeness of this study and whether this can be extrapolated to 'national' patterns and results.

Currently there is no statistic to test the geographic spread of data (rather than clustering) within a non-homogeneous population (typically residential patterns) to identify, for example, if data is falling within southern English towns rather than northern Scottish towns. Mapping all of our player data allowed us to check for obvious gaps in the national pattern, quantified by regional breakdowns.

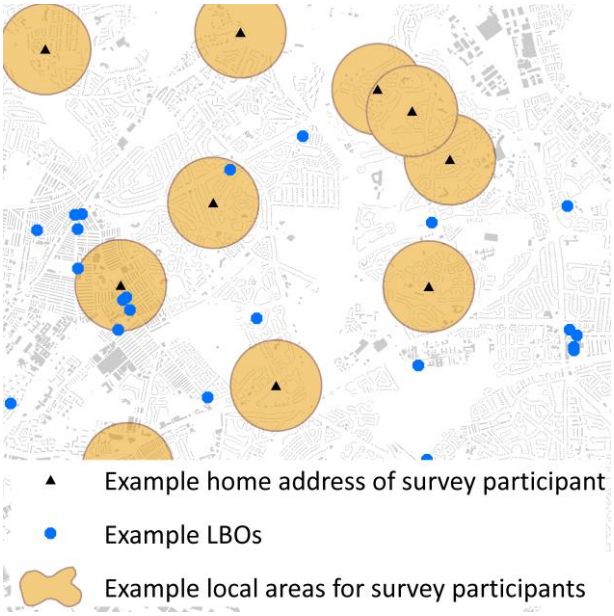
We have also briefly examined the types of neighbourhoods in which players live to provide some additional background context. We have examined geodemographic groups and the classification of urban or rural areas to briefly describe the player residences being represented. Results are discussed in section 4.

3.3 Local B2 machine counts

This analysis aims to answer the question ***how do patterns of gambling play relate to the number of B2 machine venues near a player residence?***

We have counted the number of B2 machine venues falling within 400m of each LCS participant's home address, and tested for correlations between the number of LBOs and problem gambling prevalence, number of machine play sessions and number of days machines were played. We can examine how behaviour varies based on whether someone has higher or lower numbers of LBOs accessible within their local area using the Pearson's correlation coefficient, which measures the strength of the association between two variables, and statistical variance tests.

Figure 2: Illustrative player proximate areas for analysis.



We have used 400m as the distance by which to define proximate or nearby areas to players' residences, based on the results and rationale used in our previous study (Astbury & Thurstain-Goodwin, 2015). We recognise a limitation with this approach is that 'proximity' is a subjective variable, where the results may be different depending on which distance is chosen.

In this approach we are not measuring the spatial distribution of LBO machine venues within the radius of someone's home address, only the total number available. The results can be considered a broad indication of the 'local level of provision'.

3.4 Proximity to B2 machine concentrations

This analysis aims to answer the question ***how do patterns of gambling play relate to living near a concentration of B2 machine venues?***

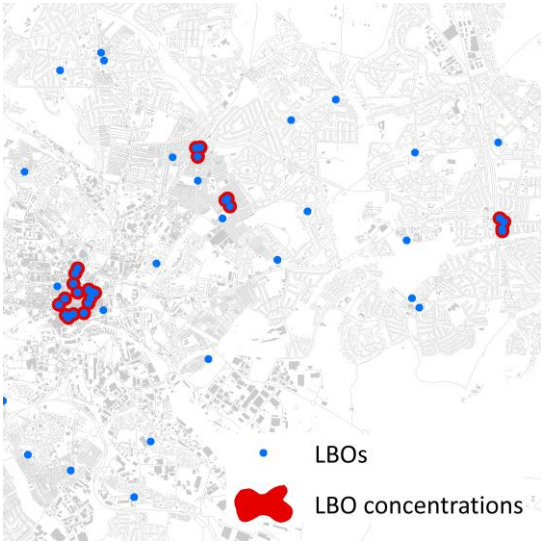
Firstly we have defined and modelled 'concentrations' (also often referred to as 'clusters') of LBOs with B2 machines. The analysis aims to capture groups of LBOs within walking-distance of one-another. We have captured groups of venues where there are 3 or more adjacent venues that are no more than either 200m or 400m away from another venue. We have thus created two sets of 'LBO concentration' areas for comparison. Those where venues are no more than 200m away from each other and those where venues are no more than 400m away from each

other. Having identified these concentrations of LBOs, we have then added a radius of 400m around them to identify the local area that these concentrations serve.

LBOs are located by unit postcode centroids (the centre of around 15 adjacent addresses), so a small error margin exists in using distances between the centroids of postcodes. However, LBOs tend to be located in urban areas where unit postcode areas are smaller in size, and we estimate this method will accurately depict concentrations well.

We have used Euclidean or straight-line distances between LBOs which cannot account for real-life obstacles and urban design.

Figure 3: Illustrative LBO concentration areas.



We have calculated the average PGSI score, problem gambling prevalence, number of machine sessions, and number of different days gambled on B2 machines for people living within and outside of these concentrations.

We have also calculated the **proportion of players** falling within and outside of these concentrations by their PGSI score, number of machine sessions, and number of different days someone gambled on B2 machines. A Pearson's Chi Squared test for independence has been used to evaluate how likely it is that any differences between the counts arose by chance.

Statistics can indicate correlation, but they cannot determine causation. For example, our statistics may show a quantitative relationship between problem gambling and living near a concentration of LBOs with B2 machines, but they cannot say that concentrations of LBOs with B2 machines are *causing* problem gambling. Indeed there often exist multiple related causal factors for further investigation.

3.5 Survey weightings

In all survey analysis, the data have been weighted to account for both non-response to the loyalty card survey and over-sampling those who were more engaged gamblers. See Wardle et al, 2014 for further details of the weighting strategy.

4. Results and findings

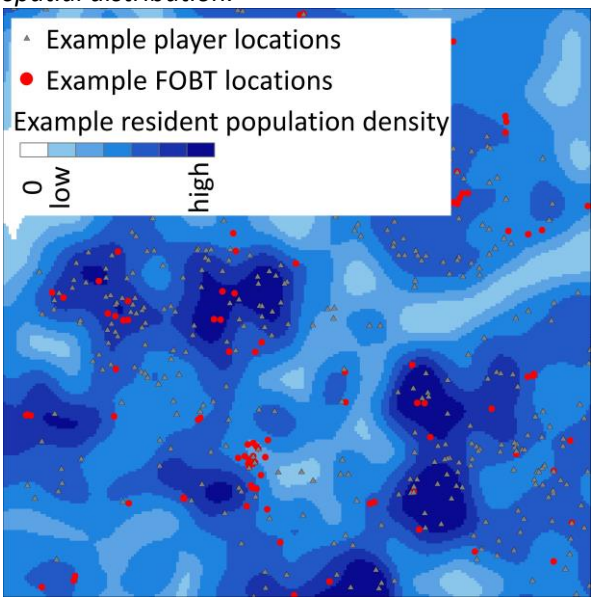
Below we discuss the results and findings from the analysis.

4.1 Profile of the sample

Clustering

For many spatial analyses we look to see if the data are randomly distributed across a study area. LBOs, by their nature, are clustered in populated areas across Great Britain. Therefore, we are not looking for complete dispersal across the country (in statistical terms referred to as complete spatial randomness or CSR). Instead we are interested in whether LBOs and players are clustered *within* these populated areas. We used the Ripley's K test to examine this.

Figure 4: Illustrative resident population and LBO spatial distribution.



Ripley's K tests show the differences in geographic clustering within individual groups of locations. The 'observed K' shows the points we are testing (e.g., LBO locations). Each dataset is tested against a randomly generated set of points (the 'expected K') representing a random distribution of points across the study area. These tests are repeated at multiple scales, or distances from each point, to see whether clustering is occurring in very local areas, or in this analysis up to 5 kilometres away.

If the observed K value is larger than the expected K value, our points are more *clustered* than a random distribution at that distance/scale. If the observed K value is larger than the 'high confidence envelope' value, spatial clustering is *statistically significant.*

If the observed K value is smaller than the expected K, our points are more *dispersed* than a random distribution at that distance/scale. If the observed K value is smaller than the 'low confidence envelope' value, spatial dispersion is *statistically significant*.

Figures 5 and 7 show that both the LBO locations and survey participants' residences are clustered at all spatial scales across the country and that this spatial clustering is statistically significant as it is much greater than the high confidence envelope for each.

We have compared the level of clustering for LBO locations and loyalty card survey participants' residences against the general distribution of residents by comparing to a random sample of

Output Area population-weighted centroids[3] across the country. If L is greater than 0, the data is indicative of being clustered; if L = 0 the data is indicative of spatially random, and if L is less than 0, the data is indicative of a dispersed distribution.

Comparing Figure 5 (LBO locations) with Figure 6 (general population distribution) we can see the value of L is higher at all distances/scales for LBO locations than the general population distribution locations. **LBOs are showing a more clustered distribution than residents.**

Comparing Figure 7 (loyalty card survey respondents' locations) with Figure 8 (general population distribution) we can see the value of L is higher at all distances/scales for survey participants than general population distribution locations. **Survey respondent residences are showing a more clustered distribution than residents.**

3 Output Areas are the smallest geographic area used to collect British Census data, representing, on average around 300 people in England and Wales and 114 people in Scotland, in 2011. Population weighted centroids are the weighted centre of an area towards where most people live.

Figure 5: Multi-variate Ripley's K statistic results for LBO locations in Great Britain.

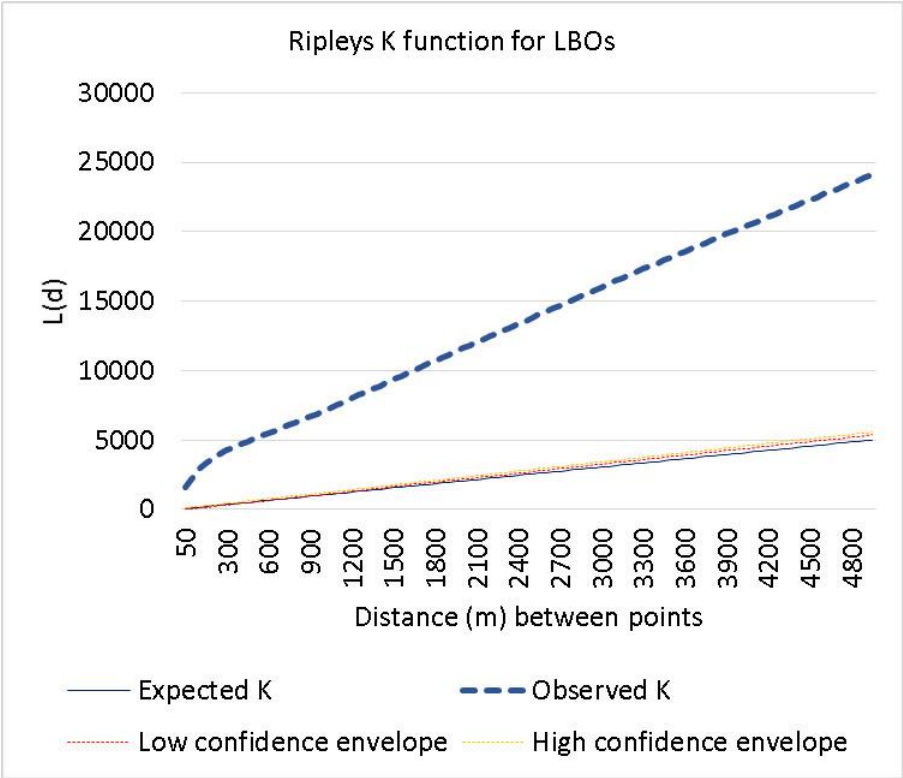


Figure 6: Multi-variate Ripley's K statistic results for sampled resident population distribution in Great Britain.

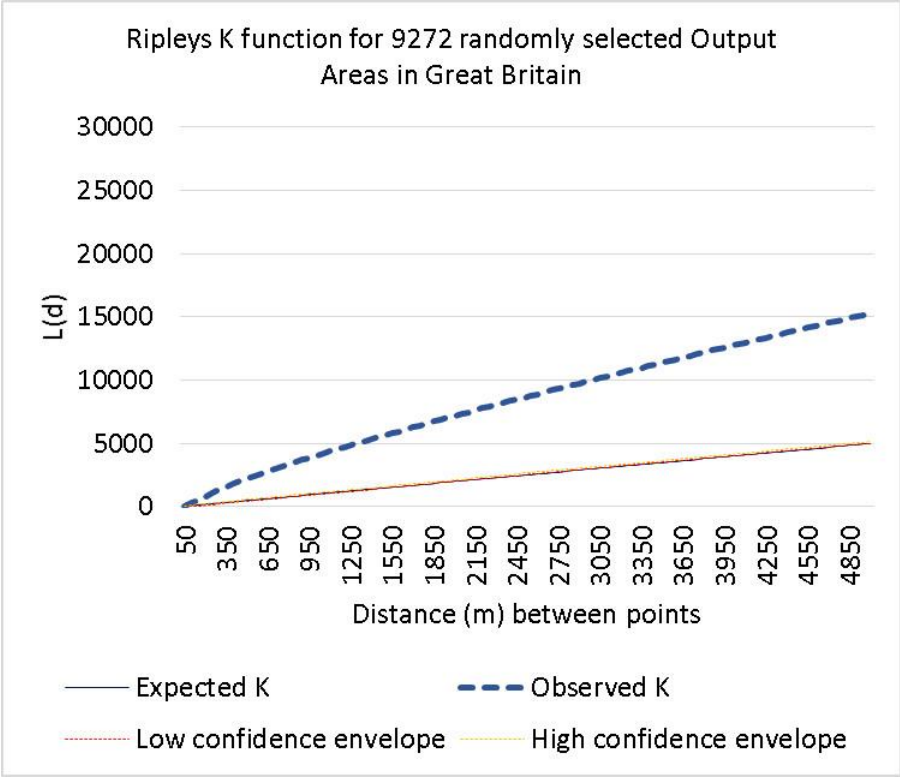


Figure 7: Multi-variate Ripley's K statistic results for survey respondent residences in Great Britain.

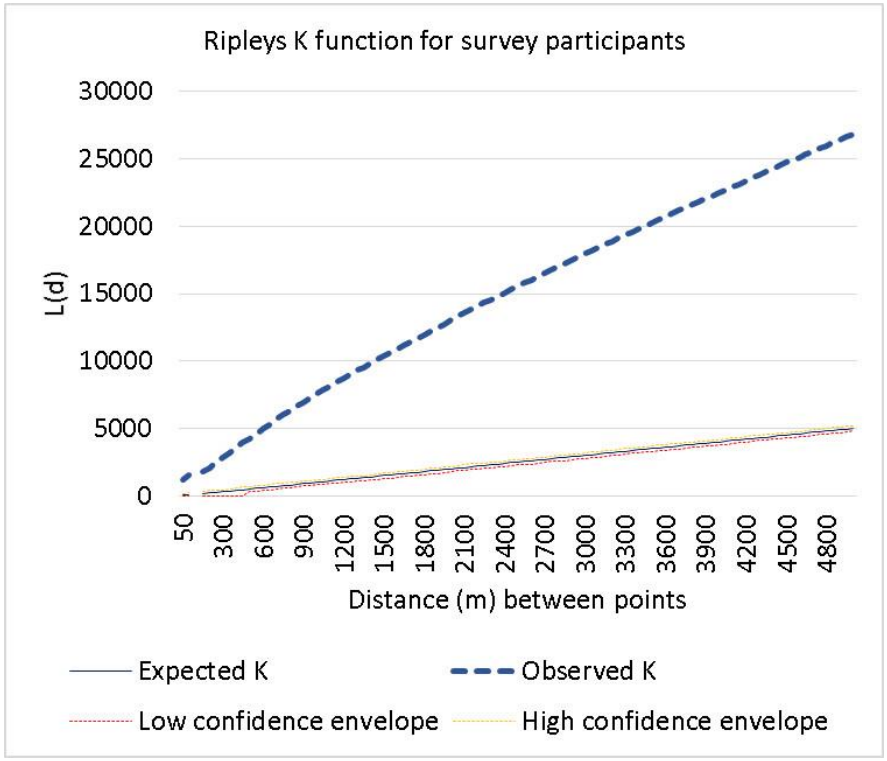
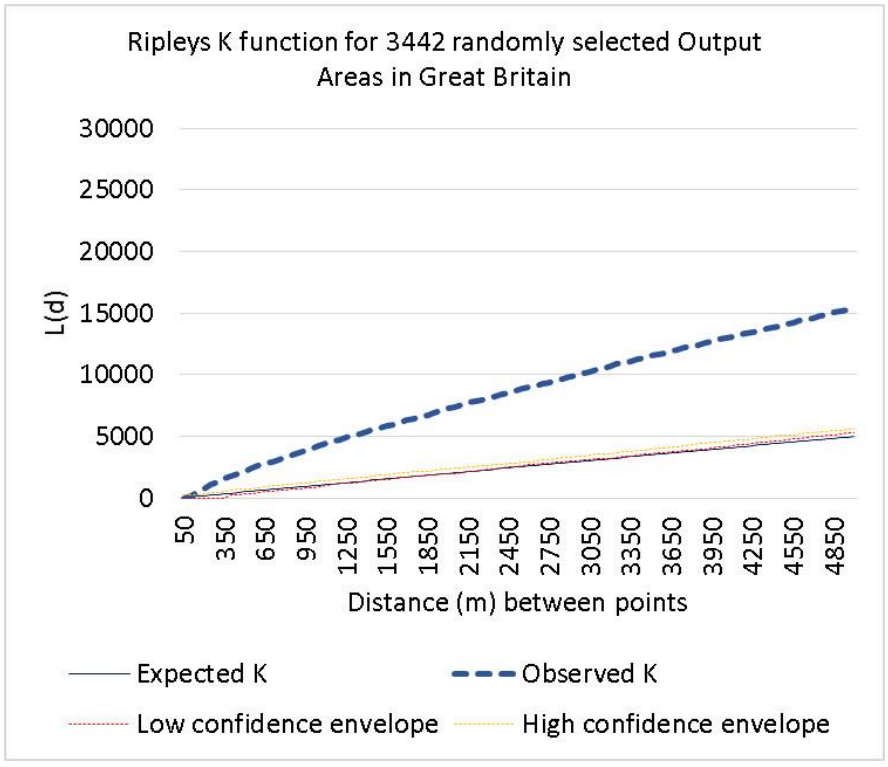


Figure 8: Multi-variate Ripley's K statistic results for sampled resident population distribution in Great Britain.



Geographic spread/representativeness

From Figures 9 and 10 we can see that our player data samples are spread across the country in a similar pattern to our population distribution. Figure 11 compares the sampled data with general population by region. This shows small regional biases, especially around London and the North East of England. For example, only around 14% of residents are in London, but among loyalty card survey participants, 18% lived in London and among players from our machine play data, 20% lived in London. Overall however the data used in this study provides a good geographic spread and representativeness across the country.

Figure 9: distribution and density of LCS participant residences in Great Britain.

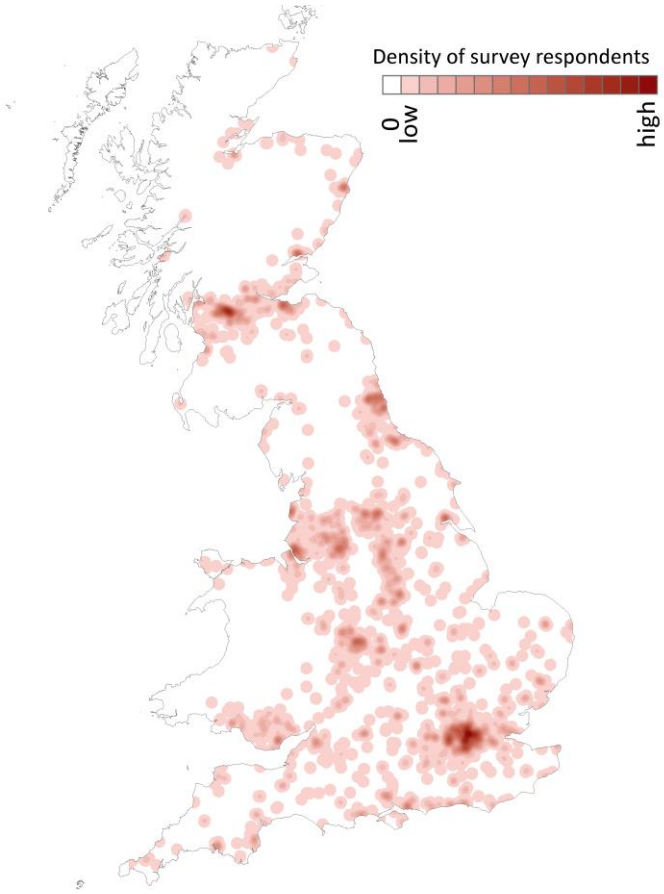


Figure 10: distribution and density of machine play data in Great Britain.

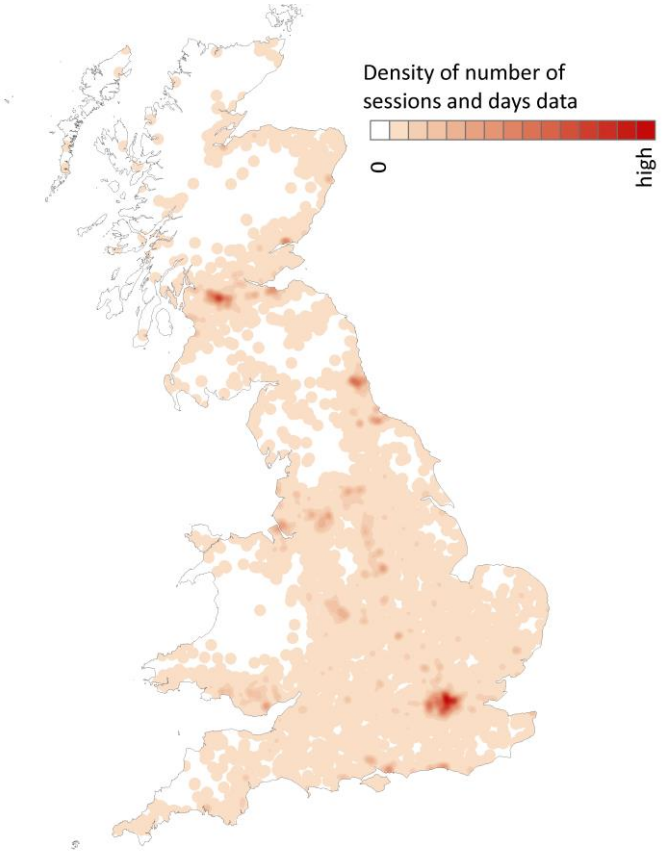
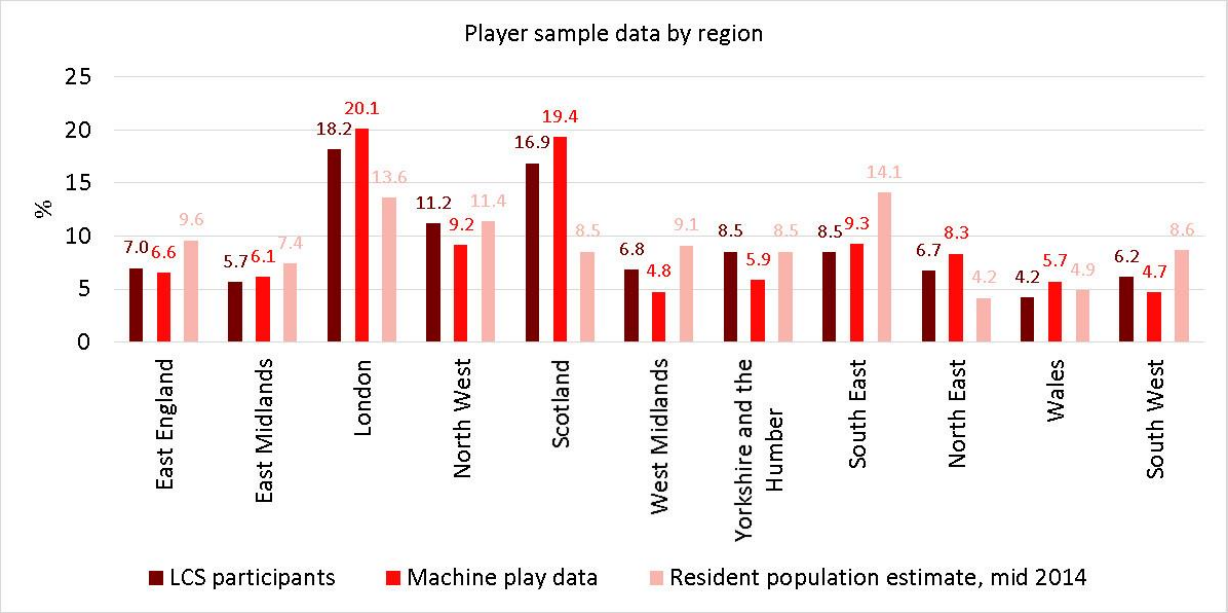


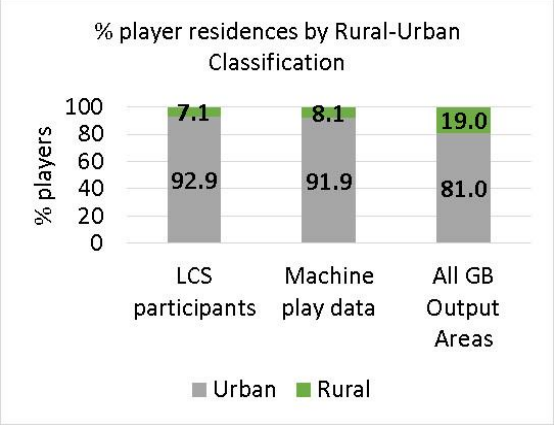
Figure 11: proportion of player sample against general population distribution by Great Britain regions. Source: LCS participant data / machine play data / Office for National Statistics / National Records of Scotland.



Neighbourhood profile

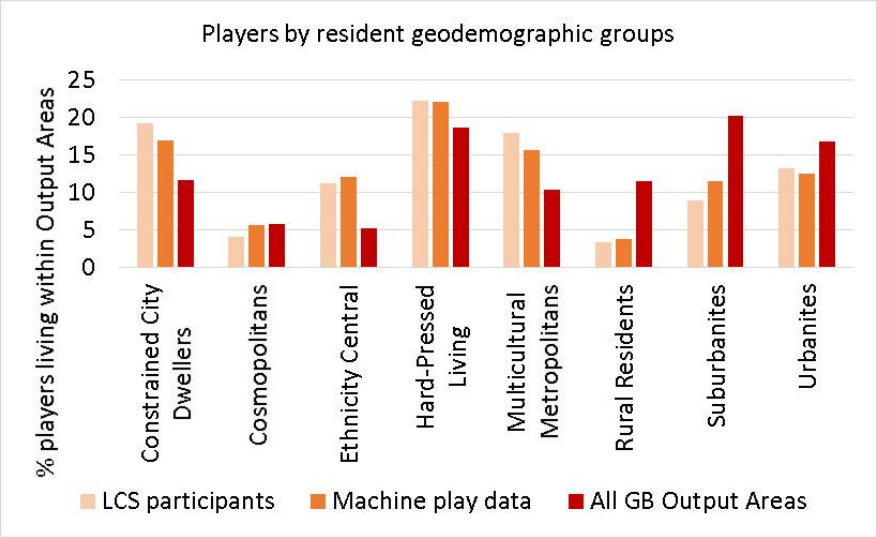
Figure 12 shows the rural or urban character of player residences (both LCS participants and people identified in the machine play data) using the Rural-Urban Classification of Output Areas, 2011[4]. Player resident postcodes are located within each corresponding Output Area. The results show our player samples to be resident in more urban areas than Great Britain (GB) on the whole. This is expanded in Figure 13, with players being over-represented in urban areas characterised as 'constrained city dwellers', 'hard-pressed living' and with multiple ethnic groups.

Figure 12: player residences by the Rural-Urban Classification, 2011.



[4] Office for National Statistics (ONS) (2013), *The 2011 Rural - Urban Classification for Small Area Geographies: A User Guide and Frequently Asked Questions (v1.0).*

Figure 13: player residences by geodemographic groups, using the Output Area Classification (OAC), 2011.[5]



5 Office for National Statistics (ONS) (2015), *Methodology Note for the 2011 Area Classification for Output Areas.*

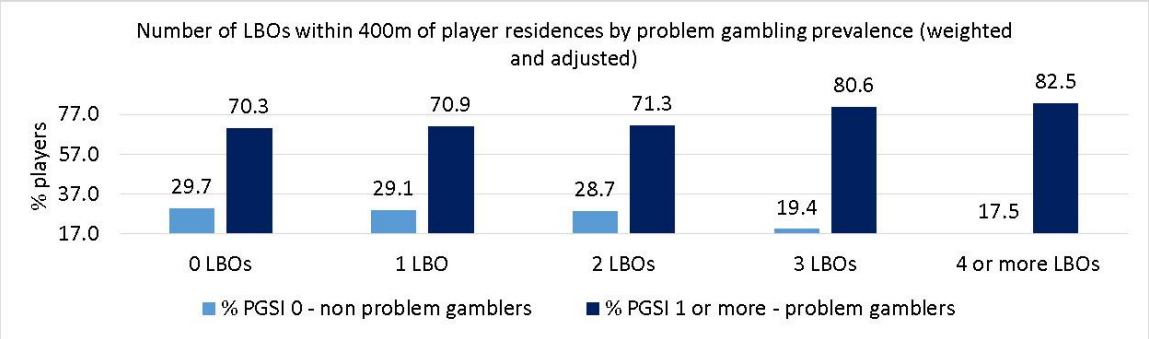
4.2 Local B2 machine counts

These results examine the relationship between LBO counts within 400m of player home addresses and gambling behaviour.

Problem gambling

Figure 14 below shows the distribution of PGSI scores by the number of nearby LBOs.

Figure 14: Number of LCS participants by the number of nearby LBOs and PGSI score.



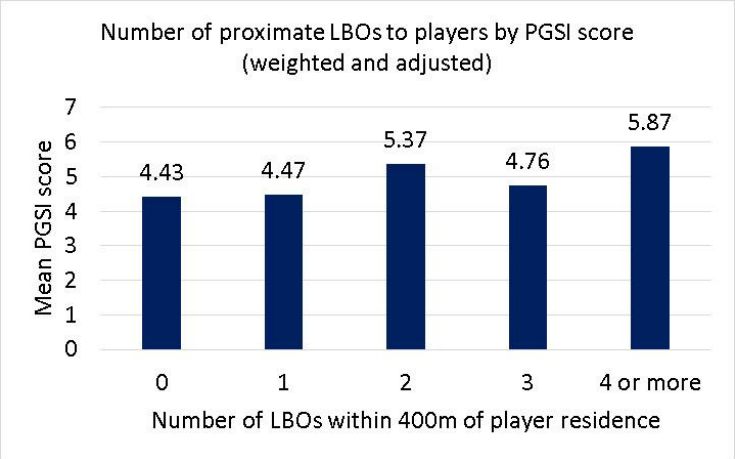
There were no statistically significant differences in problem gambling prevalence according to the number of LBOs someone had within 400m of the survey participant's home when using the full range of PGSI scores. However, when combining rates of low risk, moderate risk and problem gambling (a PGSI score of 1 or more), a statistically significant relationship was detected.

Around 30% of those who had no LBOs in their local area were non-problem gamblers compared with 18% among those who had four or more LBOs in their local area, (Figure 14, Appendix Table A). This means that the prevalence of having a PGSI score of 1 or more (i.e., at least a low risk gambler) was higher among those living in areas with 3 or 4 LBOs in their immediate location.

We have also examined how the mean PGSI scores of LCS participants varied by how many LBOs were nearby. The results are shown in Figure 15.[6] Average PGSI scores did not vary statistically significantly according to the number of LBOs in the immediate local area of LCS participants (Appendix Table B). These results may, in part, be due to small base sizes since only 142 LCS participants lived in areas with four or more LBOs in their local area, making detecting a significant difference between groups difficult.

[6] These analyses were tested using an adjusted Wald's F test in SPSS v19 which takes into account the complex survey design and weighting for the survey data.

Figure 15: Mean PGSI scores by the number of LBOs within 400m of a LCS participant residence (weighted and adjusted).



Number of sessions played

A Pearson's correlation test was performed to examine if there was any relationship between the number of sessions gambled on B2 machines and the number of LBOs within 400m of a player's residence. There was no statistically significant linear relationship shown (r = -0.0048, df = 179,041, P <.005). We cannot say that the number of sessions either increased or decreased according to the number of LBOs in a player's local area.

Number of different days played

A Pearson's correlation test was performed to examine if there are any relationship between the number of days played and the number of LBOs within 400m of a player's residence. There was no statistically significant linear relationship shown (r = -0.0209, df = 179,041, P <.005). We cannot say that the number of days gambled either increase or decrease according to the number of LBOs in a player's local area.

4.3 Proximity to B2 machine concentrations

These results examine the relationship between LBO concentrations close to players' homes and gambling behaviour. We have calculated two measures of 'B2 concentrations' – those where venues are within 200m of each other and those where venues are within 400m of each other.

Problem gambling

Figures 16 and 17 below show the proportion of players living within LBO concentrations by PGSI scores. Overall, the prevalence of problem gambling was higher among those who lived within a 400m LBO concentration than those who did not. 28.1% of those who lived within a 400m LBO concentration area were problem gamblers compared with 22.1% of those who did not. Likewise, mean PGSI scores were significantly higher among those who lived within a 400m concentration area (5.4) than those who did not (4.5) (see Table 2).

Similar patterns were observed when looking at 200m LBO concentration areas. Mean PGSI scores were higher among those who lived within a 200m LBO concentration (5.5) than those who did not (4.5). Rates of moderate risk and problem gambling were higher among those living in 200m LBO locations (58%) than those who did not (46%). Problem gambling rates were 28% and 22% respectively (the same as 400m LBO concentrations). However, because fewer people lived in 200m concentrations, this was not statistically significant (the p value was 0.099; Appendix Table C). This is simply a function of sample sizes, making differences more difficult to detect. All other results were statistically significant (Appendix Tables D, E and F).

Figures 18 and 19 show the distribution of PGSI scores according to whether someone lived in a 400m LBO concentration area or not, and whether someone lived in a 200m LBO concentration area or not. Looking at Figure 19 (distribution of PGSI score according to whether someone lives in a 400m LBO concentration area or not) we can see that up to the 35th centile, PGSI scores are broadly similar between those who do and who do not live in high concentration areas. From the 40th centile onwards, PGSI scores tend to diverge between the two groups and are higher among those living in LBO concentration areas. For example, 40% of those living in high LBO concentration areas have a PGSI score of 5 or more (the 60th centile). PGSI scores only reach this level at the 70th centile for those not living in these areas. At the 95th percentile for those living in high LBO concentration areas, PGSI scores are 20 whereas among those who do not live in these areas, PGSI scores at the 95th percentile are 18.

Figure 16: The proportion of players by PGSI score and whether they lived in a 200m LBO concentration.

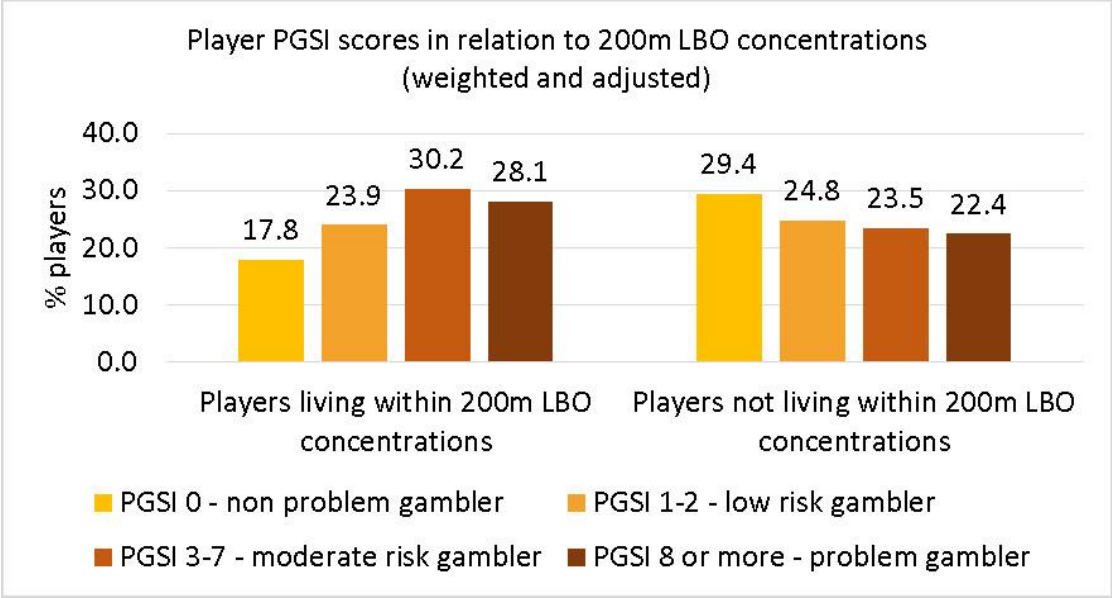


Figure 17: The proportion of players by PGSI score and whether they lived in a 400m LBO concentration.

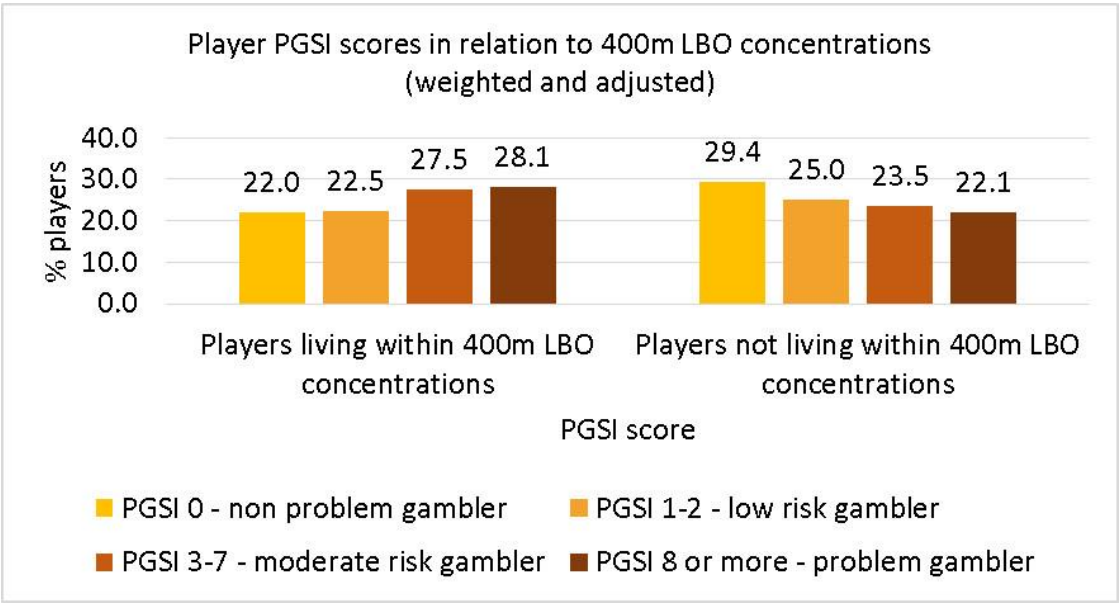


Table 2: Mean PGSI scores (weighted and adjusted) by the residence of players in relation to LBO concentrations.

	Lives within an LBO concentration	*Does not live within an LBO concentration*
200m LBO concentrations	**5.51**	**4.52**
400m LBO concentrations	**5.42**	**4.48**

Figure 18: Frequency distribution of PGSI scores in relation to 200m LBO concentrations.

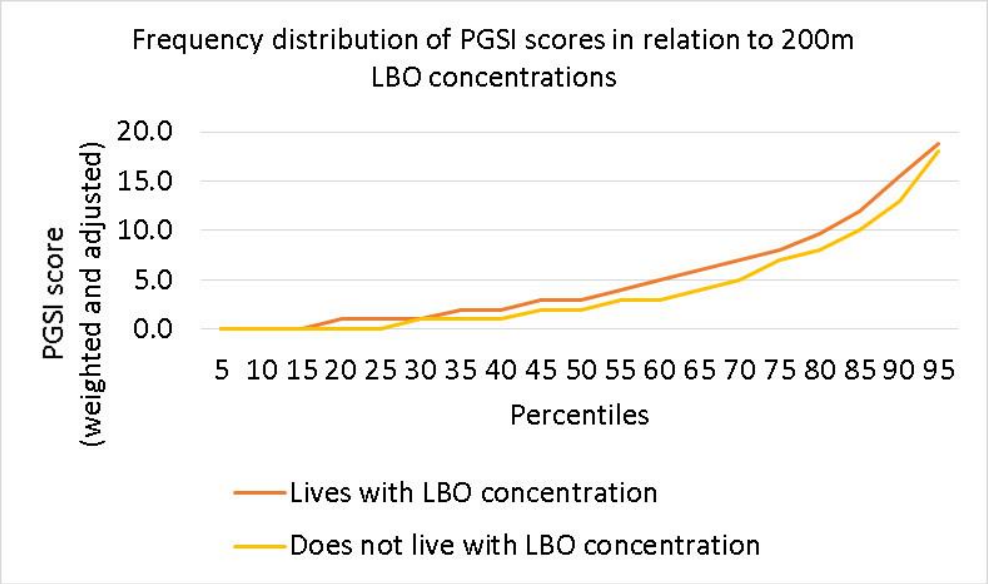
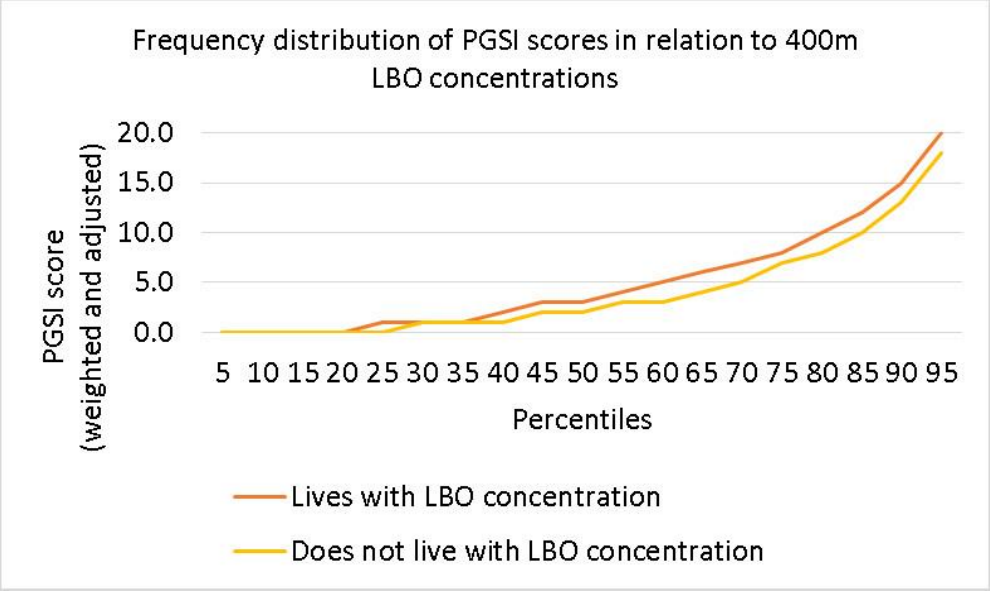


Figure 19: Frequency distribution of PGSI scores in relation to 400m LBO concentrations.



Number of sessions played

Figures 20 and 21 show the trends between the proportions of players living in LBO concentrations and the number of sessions played on B2 machines, using machine play data from operators. There is a statistically significant relationship showing that players living within LBO concentrations play slightly fewer sessions (Appendix Tables I and J). This suggests that whilst PGSI scores tend to be higher in these concentrations, the number of sessions played is marginally smaller.

Table 3 shows the average number of machine play sessions among those who lived in an LBO concentration and those who did not. Looking at both 200m and 400m concentrations, the average number of machine play sessions were lower among those who lived within a LBO concentration than those who did not. The average number of B2 machine sessions was 168.9 for those who lived within a 400m LBO concentration and 229.2 for those who did not. The same pattern was observed for 200m LBO concentrations where mean numbers of B2 sessions were 172.3 vs 225 respectively (Appendix Tables M and N).

Figure 20: Proportion of players by number of B2 sessions and whether they lived within a 200m LBO concentration.

Figure 21: Proportion of players by number of B2 sessions and whether they lived within a 400m LBO concentration.

Table 3: Mean number of sessions per unique shop visited per player.

	Lives within an LBO concentration	*Does not live within an LBO concentration*
200m LBO concentrations	**172.3**	**225**
400m LBO concentrations	**168.9**	**229.2**

Number of different days played

We also analysed the number of days spent gambling on B2 machines by LBO concentrations, and the results were similar (Figures 22 and 23). The pattern suggests that the number of days played is slightly smaller when living within LBO concentrations. The average number of gambling days on B2 machines was lower among those who lived within LBO concentrations (6.8 days) than those who did not (7.9); see Table 4. These results were statistically significant (Appendix Tables G, H, K and L).

Figure 22: Proportion of players by number of days gambled and whether they lived within a 200m LBO concentration.

Figure 23: Proportion of players by number of days gambled and whether they lived within a 400m LBO concentration.

Table 4: Mean number of days played per unique shop visited per player.

	Lives within an LBO concentration	*Does not live within an LBO concentration*
200m LBO concentrations	**6.8**	**7.9**
400m LBO concentrations	**6.8**	**7.9**

4.4 Known error margins and limitations

Sampling

All results are sampled from loyalty card records, which is a sample of highly-engaged players. The results should be viewed as representative of players holding loyalty cards, rather than all players in Great Britain.

Furthermore, we know that not all loyalty card holders use their card every time they gamble. This can affect data on total counts of sessions or number of days gambled, as used in this study. There may be some systematic bias by which those who live in LBO concentrations use their loyalty cards less often than those who do not. This could affect the results shown in this report.

Because of the size of data samples, locations will be widely dispersed across Great Britain, with no one local area containing enough information to map local trends. Because of this our models are limited to global (overall) statistics about the whole country but we cannot examine local variation.

Attributes

The calculation of the problem gambling index score (PGSI) has been validated on a Canadian population only. There have been no tests to ascertain whether the score translates to a British context, and whilst we are confident this score gives a valid indication of problem gambling, this methodology should ideally be further tested in context.

Modelling

Our engagement statistics include the number of sessions gambled per player per shop, which is averaged over the time period of data collection. These averages struggle to pick out potential episodic gambling which can vary over time, where a large number of sessions may be played in a single day.

For accessibility we have used the Euclidean (straight-line) distance rather than real-life accessibility to LBOs.

Players are geo-located by the postcode of their resident location, which was the only geographic identifier within the available information. The analysis is therefore limited to residential patterns, and does not capture any patterns reflected in the 'daytime' locations of players.

Finally, due to data availability and sample sizes, we have not identified the clusters in which players are gambling in full. Players may not necessarily be gambling within the local LBOs and LBO concentrations which we have attributed to a player.

4.5 Suggested extensions to research

Several extensions to the research may help to expand on our understanding of these results. The following analysis is suggested;

- Measuring more individual frequency of play or engagement metrics.
- Sensitivity testing different sizes of LBO clusters, with a larger minimum number of LBOs or walking distance between LBOs.

- Incorporating the number of LBOs in the cluster into the model to account for the size of LBO clusters.
- Segmenting the results to examine the impact of the type of location in which the LBO cluster or player is located. This may include in or out-of-town centre locations, or the types or size of town or cities.
- Segmenting the results to see if the socio-economic and demographic player characteristics have an effect on the results. This may include age, gender, ethnicity, economic activity, or other pertinent characteristics as identified in previous gambling studies.
- Focusing on players who have definitely played in the nearby clusters identified. This is likely to require a bigger survey sample size to achieve a statistically significant result.
- Identifying how the extent of machine play specifically contributes to problem gambling.

5. Conclusions and key findings

This study has provided an opportunity to use a unique set of national data on gambling behaviour which can be mapped using the home locations of players and compared with the locations of LBOs. It has allowed us to see if the spatial distribution of LBOs with B2 machines, and living near certain LBO spatial configurations has any relationship with problem and frequency of gambling and B2 machine play in Great Britain.

We have identified that within Great Britain, LBOs cluster within populated areas. From the spatial distribution of LBOs we have been able to define higher-concentration areas using two different sizes for comparison. The resident location of players has been compared to both the proximity of one of these LBO concentrations, as well as the simple count of LBOs near to players' homes.

To do this, we looked at whether problem gambling and B2 machine play varied according to whether someone lived in an area with a greater or lesser number of LBOs and whether they lived in an area with LBO concentrations or not. LBO concentrations meant that at least three LBOs were within either 400m or 200m of one another.

Our results suggest a relationship between the spatial distribution of LBOs and problem gambling. **Mean PGSI scores and problem gambling prevalence rates were higher among those living in either LBO concentration areas than those who did not**. When looking at the relationship by simply counting the number of LBOs within a player's locale, we saw that gambling prevalence rates were higher when combining rates of low risk, moderate risk and problem gambling (a PGSI score of 1 or more), compared to non-problem gamblers, where there are more LBOs in the local area. However, the relationship with problem gambling was clearer and stronger when looking at concentrations of LBOs, suggesting that it is the **concentration of LBOs rather than the absolute number that may be important** in this relationship.

This analysis shows a strong association between concentrations of LBOs and problem gambling prevalence rates but we also need to recognise that **other factors could be driving this**. For example, concentrations of LBOs may be disproportionately in deprived areas, which are known to be associated with higher problem gambling rates. This underlying factor could, therefore, be driving this relationship. This analysis also measures problem gambling generally across all forms of activity and we do not know the level of specific problems associated with machines. Regardless of the direction of causation or other potential explanations, the spatial pattern is important. This is the first time it has been demonstrated that, for whatever reason, rates of problem gambling are higher among those who live in LBO concentrations and suggests that those living within LBO concentrations may be considered an at-risk or vulnerable group.

When looking specifically at the machine play behaviour, our results showed that **living near to an LBO concentration was associated with slightly lower frequency of play, on average**. We only looked at two machine behaviour variables: total number of B2 machine sessions and total number of gambling days, since these were the only data available. Notably, these two variables may be affected by biases in how often people use their loyalty cards when playing machines, which may have affected results. However, the extent to which this may have occurred is unknown.

The results should be interpreted within the context of the data. There were several limitations to the analysis worth noting when interpreting the results. First, the extent to which people gambled in the LBOs close to their homes is unknown. Second, we were only able to examine

the geographic location of players by their home residence. People are often found in a different 'daytime' location, which may be a set of multiple, equally important locations from which players are accessing machine play.

Finally, the findings are not representative of all machine players, but rather a subset of highly engaged gamblers who had a loyalty card for one of three bookmakers in 2014. The extent to which results can be generalised to the broader population of machine players is unknown.

Despite these limitations, this is the first time a spatial relationship between problem gambling rates and proximity to LBO concentrations has been demonstrated in Great Britain. This provides important new insight and suggests that further work should be undertaken to explore and understand the mechanisms which underpin this relationship.

References

Astbury, G. & Thurstain-Goodwin, M. (2015) *Contextualising machine gambling characteristics by location - A spatial investigation of machines in bookmakers using industry data.* Geofutures Ltd, Bath. http://responsiblegamblingtrust.org.uk/user_uploads/b2gamingmachines_finalreport_20150218.pdf Accessed January 2016.

Office for National Statistics (ONS) (2013) *The 2011 Rural - Urban Classification for Small Area Geographies: A User Guide and Frequently Asked Questions (v1.0).* https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/239478/RUC11user_guide_28_Aug.pdf Accessed January 2016.

Office for National Statistics (ONS) (2015) *Methodology Note for the 2011 Area Classification for Output Areas.* http://www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/ns-area-classifications/ns-2011-area-classifications/methodology-and-variables/index.html Accessed January 2016.

Wardle, H., Keily, R., Astbury, G. & Reith, G. (2012) *'Risky Places?': Mapping Gambling Machine Density and Socio-Economic Deprivation*, Journal of Gambling Studies, Vol 30 (1), pp 201-212.

Wardle, H., Parke, J. & Excell, D. (2014a) *Machines Research Programme: Report 1 - Theoretical markers of harm for machine play in a bookmaker's. A rapid scoping review*. NatCen Social Research, London. http://www.responsiblegamblingtrust.org.uk/user_uploads/pdfs/report%201%20theoretical%20markers%20of%20harm%20for%20machine%20play%20in%20a%20bookmakers%20-%20a%20rapid%20scoping%20review.pdf Accessed January 2016.

Wardle, H., Excell, D., Ireland, E., llic, N. & Sharman, S. (2014b) *Gambling Machines Research Programme: Report 2 - Identifying problem gambling – findings from a survey of loyalty card customers*. Natcen Social Research, London. http://www.responsiblegamblingtrust.org.uk/user_uploads/pdfs/report%202%20identifying%20problem%20gambling%20-%20findings%20from%20a%20survey%20of%20loyalty%20card%20customers.pdf Accessed January 2016.

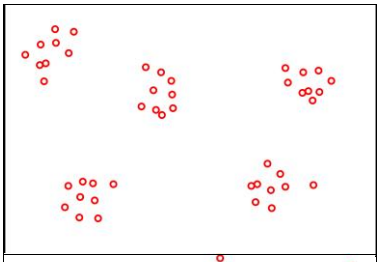
Wardle, H., Ireland, E., Sharman, S., Excell, D. & Gonzalez-Ordonez, D. (2014c) *Patterns of play: analysis of data from machines in bookmakers.* Natcen Social Research, London. http://www.responsiblegamblingtrust.org.uk/user_uploads/pdfs/patterns%20of%20play%20-%20analysis%20of%20data%20from%20machines%20in%20bookmakers.pdf Accessed January 2016.

Technical Appendix A

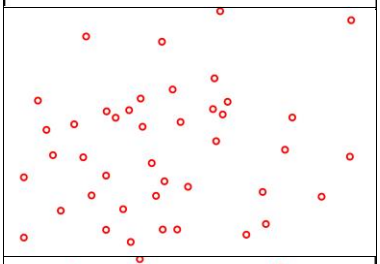
Ripley's K statistic

The Ripley's K statistic measures the degree of clustering by calculating the 'intensity' of points across a study area, using circles of progressively larger size. A circle is placed over a point, the number of points is counted within the circle, the circle moves to every other point in turn counting points. If the average number of points is higher than the average intensity in the whole study area, the distribution is considered clustered. This is calculated for each distance circle to show spatial clustering at different neighbourhood sizes.

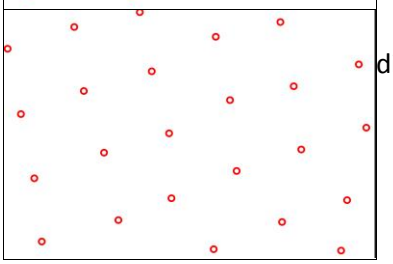
Results are transformed to make them linear (L), and plotted against distance to display clustering at different spatial scales. We are looking for clustering within the already-clustered overall population distribution; to do this we can compare the results from the distribution of residents against the distribution of LBOs to look for differences in cluster patterns.



If L is greater than 0, the data is indicative of clustered;



If L = 0 the data is indicative of spatially random;



and if L is less than 0, the data is indicative of a dispersed distribution.

If the observed K value is larger than the expected K value, our points are more *clustered* than a random distribution at that distance (scale of analysis). If the observed K value is larger than the 'high confidence envelope' value, spatial clustering is *statistically significant.*

If the observed K value is smaller than the expected K, our points are more *dispersed* than a random distribution at that distance. If the observed K value is smaller than the 'low confidence envelope' value, spatial dispersion is *statistically significant*.

Results Appendix B

Detailed results from statistical tests carried out for this analysis.

Wald's F tests

These numbers are weighted and adjusted to account for oversampling for problem gamblers.

Table A: Proportions of players by PGSI scores and the prevalence of LBOs near to players.

PGSI Score	*Number of LBOs within 400m of player residence*					Totals
	0	*1*	*2*	*3*	*4 or more*	
0 - non problem gambler	641	207	92	28	25	993
1-2 - low risk gambler	536	185	64	44	32	861
3-7 - moderate risk gambler	518	158	76	35	47	834
8 or more - problem gambler	460	162	89	37	39	787
Totals	2155	712	321	144	143	3475

Problem gamblers

Source	*df1*	*df2*	*Wald F*	*Sig.*
(Corrected Model)	4.000	3438.000	1.293	.270

Non problem-gamblers

Source	*df1*	*df2*	*Wald F*	*Sig.*
(Corrected Model)	4.000	3438.000	2.375	.050

Table B: Mean PGSI scores by the prevalence of LBOs near to players.

Number of LBOs within 400m of player residence	*Mean PGSI score estimate*	*Standard Error*	*Design Effect*	*Square Root Design Effect*
0	4.43	.171	1.861	1.364
1	4.47	.285	1.707	1.306
2	5.37	.495	1.836	1.355
3	4.76	.548	1.503	1.226
4	5.87	.683	1.702	1.305

Source	*df1*	*df2*	*Wald F*	*Sig.*
(Corrected Model)	4.000	3436.000	1.775	.131

Table C: Proportions of players by PGSI scores and the location of 200m LBO concentrations in relation to players (weighted and adjusted).

	Lives within a 400m LBO concentration	*Does not live within a 400m LBO concentration*	Total
0 - non problem gambler	50	940	990
1-2 - low risk gambler	67	791	858
3-7 - moderate risk gambler	84	750	834
8 or more - problem gambler	78	715	793
Totals	279	3196	3475

Problem gamblers alone

Source	df1	df2	Wald F	Sig.
(Corrected Model)	1	3441	2.723	.099

Moderate risk and problem gamblers only

Source	df1	df2	Wald F	Sig.
(Corrected Model)	1	3441	9.255	.002

Non problem gamblers

Source	df1	df2	Wald F	Sig.
(Corrected Model)	1	3441	11.337	.001

Table D: Proportions of players by PGSI scores and the location of 400m LBO concentrations in relation to players (weighted and adjusted).

	Lives within a 400m LBO concentration	*Does not live within a 400m LBO concentration*	Total
0 - non problem gambler	99	891	990
1-2 - low risk gambler	101	757	858
3-7 - moderate risk gambler	123	711	834
8 or more - problem gambler	126	667	793
Totals	449	3026	3475

Source	df1	df2	Wald F	Sig.
(Corrected Model)	1.000	3441.000	6.567	.010

Table E: Mean PGSI scores by the location of 200m LBO concentrations in relation to players (weighted and adjusted).

	Mean PGSI score estimate	*Standard Error*	*Design Effect*	*Square Root Design Effect*
Players living within a 200m LBO concentration	5.51	.443	1.567	1.252
Players not living within a 200m LBO concentration	4.52	.141	1.818	1.348

Source	*df1*	*df2*	*Wald F*	*Sig.*
(Corrected Model)	1	3439	4.524	.033

Table F: Mean PGSI scores by the location of 400m LBO concentrations in relation to players (weighted and adjusted).

	Mean PGSI score estimate	*Standard Error*	*Design Effect*	*Square Root Design Effect*
Players living within a 400m LBO concentration	5.42	.359	1.528	1.236
Players not living within a 400m LBO concentration	4.48	.145	1.846	1.359

Source	*df1*	*df2*	*Wald F*	*Sig.*
(Corrected Model)	1.000	3439.000	5.867	.015

Chi squared independence tests

Table G: Proportions of players by the number of days played and the location of 200m LBO concentrations in relation to players.

Number of different days played per unique shop visited per player	*Players living within a 200m LBO concentration*		*Players not living within a 200m LBO concentration*		**Totals**
	Observed	*Expected*	*Observed*	*Expected*	
1	8127	7503	72873	73497	81000
2 to 5	4922	5115	50296	50103	55218
6 to 10	1314	1508	14968	14774	16282
11 to 20	1000	1080	10662	10582	11662
21 to 30	434	449	4416	4401	4850
31 to 40	199	245	2451	2405	2650
41 to 50	140	161	1602	1581	1742
51 to 100	290	337	3346	3299	3636
101 to 200	144	155	1529	1518	1673
201 or more	14	31	316	299	330
Totals	16584		162459		179043

Chi squared calculation =**130.6** p<**0.00001**

Table H: Proportions of players by the number of days played and the location of 400m LBO concentrations in relation to players.

Number of different days played per unique shop visited per player	*Players living within a 400m LBO concentration*		*Players not living within a 400m LBO concentration*		**Totals**
	Observed	*Expected*	*Observed*	*Expected*	
1	12866	12118	68134	68882	81000
2 to 5	8056	8261	47162	46957	55218
6 to 10	2217	2436	14065	13846	16282
11 to 20	1692	1745	9970	9917	11662
21 to 30	700	726	4150	4124	4850
31 to 40	333	396	2317	2254	2650
41 to 50	221	261	1521	1481	1742
51 to 100	464	544	3172	3092	3636
101 to 200	210	250	1463	1423	1673
201 or more	27	49	303	281	330
Totals	26786		152257		179043

Chi squared calculation =**138.7** p<**0.00001**

Table I: Proportions of players by the number of sessions played and the location of 200m LBO concentrations in relation to players.

Number of sessions played per unique shop visited per player	*Players living within a 200m LBO concentration*		*Players not living within a 200m LBO concentration*		**Totals**
	Observed	*Expected*	*Observed*	*Expected*	
1	5068	4378	42195	42885	47263
2 to 5	5239	5483	53952	53708	59191
6 to 10	1742	1857	18305	18190	20047
11 to 20	1328	1468	14524	14384	15852
21 to 50	1363	1395	13694	13662	15057
51 to 100	599	697	6931	6833	7530
101 to 500	818	844	8297	8271	9115
501 to 1000	166	158	1535	1543	1701
1001 to 2000	86	92	907	901	993
2001 or more	175	212	2119	2082	2294
Totals	16584		162459		179043

Chi squared calculation = **179.7** p<**0.00001**

Table J: Proportions of players by the number of sessions played and the location of 400m LBO concentrations in relation to players.

Number of sessions played per unique shop visited per player	*Players living within a 400m LBO concentration*		*Players not living within a 400m LBO concentration*		**Totals**
	Observed	*Expected*	*Observed*	*Expected*	
1	7934	7071	39329	40192	47263
2 to 5	8453	8855	50738	50336	59191
6 to 10	2877	2999	17170	17048	20047
11 to 20	2249	2372	13603	13480	15852
21 to 50	2239	2253	12818	12804	15057
51 to 100	1034	1127	6496	6403	7530
101 to 500	1326	1364	7789	7751	9115
501 to 1000	248	254	1453	1447	1701
1001 to 2000	139	149	854	844	993
2001 or more	287	343	2007	1951	2294
Totals	26786		152257		179043

Chi squared calculation = **180.7** p<**0.00001**

Students T-tests with F-tests

Students T tests for mean values with F tests for homogeneity of variance assumption at 0.05 probability.

Table K: Mean number of days played by the location of 200m LBO concentrations in relation to players.

	Variable 1	*Variable 2*
Mean	6.827122528	7.852233
Variance	306.8343099	414.2565
Observations	16584	162459
T test		
Hypothesized Mean Difference	0	
df	21427	
t Stat	-7.065238998	
P(T<=t) two-tail	1.65E-12	
t Critical two-tail	1.960074705	
F test		
F	1.35009845	
P two-tail	0	
F Critical two-tail	1.022962025	

Table L: Mean number of days played by the location of 400m LBO concentrations in relation to players.

	Variable 1	*Variable 2*
Mean	6.831442	7.920161
Variance	305.4881	421.6178
Observations	26786	152257
T test		
Hypothesized Mean Difference	0	
df	40946	
t Stat	-9.14474	
P(T<=t) two-tail	6.24E-20	
t Critical two-tail	1.960022	
F test		
F	1.380144801	
P two-tail	0	
F Critical two-tail	1.018596673	

Table M: Mean number of sessions played by the location of 200m LBO concentrations in relation to players.

	Variable 1	*Variable 2*
Mean	172.3161	225.0468
Variance	5602064	10947406
Observations	16584	162459
T test		
Hypothesized Mean Difference	0	
df	23762	
t Stat	-2.61961	
P(T<=t) two-tail	0.008809	
t Critical two-tail	1.960064	
F test		
F	1.95417383	
P two-tail	0	
F Critical two-tail	1.022962025	

Table N: Mean number of sessions played by the location of 400m LBO concentrations in relation to players.

	Variable 1	*Variable 2*
Mean	168.9045	229.1803
Variance	5576400	11309820
Observations	26786	152257

T test		
Hypothesized Mean Difference	0	
df	48229	
t Stat	-3.58641	
P(T<=t) two-tail	0.000336	
t Critical two-tail	1.960013	
F test		
F	2.028158022	
P two-tail	0	
F Critical two-tail	1.018596673	

For further information please contact:

Mark Thurstain-Goodwin
Managing Director
Geofutures Limited
39 Powlett Road
Bath BA2 6QL

Tel: 01225 222382

Email: mtg@geofutures.com
www.geofutures.com