Effects of Income on National Lottery Participation in the UK

Student ID[1]: 1720320
Word Count: 4998

Abstract

This paper draws upon the Friedman-Savage model to investigate the relationship between income and regular national lottery participation in the UK. A distinction between household income and personal income as two different income variables is made to account for the difference in income measures used in different literatures. Data from the UK gambling prevalence survey conducted in 2010 is used. Using binary probit models, a positive, reverse U-shaped relationship between both income variables and tendencies for monthly lottery participation is found in regressions that only include either personal income or household income variables as sole measures of income. In regressions that include both variables of income, a similar reverse U-shaped trend can be found for household income variables, while a positive linear relationship between personal income and tendencies for monthly lottery participation is found instead. This paper argues how effects of wealth inherent in income is captured by household income variables, while personal income variables can potentially reveal a relative measure of income satisfaction that can be captured only when household income variables are controlled.

[1] I would like to thank Prof. Alexander Dobson for his continued help and support throughout my research process. This paper would not be possible without his guidance and expertise on behavioral economics and econometrics.

Table of Contents

Introduction

In a recent report, the UK Gambling Commission (2019) reported that British gamblers have spent £7.2 billion on national lottery tickets during the year of March 2019, a £271.6m increase from the previous year. With approximately 70% of the UK's over 18s partaking in the national lottery, and 50% participating more than once a month, it is clear that lottery participation has grown to become a prevalent recreational activity among UK citizens since its introduction by the government in 1994. But what motivates individuals to consistently partake in low risk, high reward gambles? As many theories in behavioral economics such as prospect theory suggest a strong causal relationship between one's relative income and risk attitudes to gambles, I am particularly interested to observe whether similar trends can be observed under a controlled environment in gambling, where risks from gambles can be measured with more certainty under (relatively) more easily calculable probabilities.

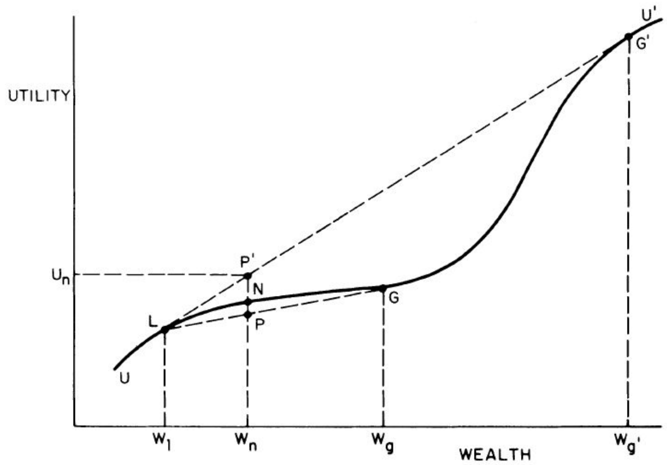
This study aims to find the relationship between income and participation in the national lottery. For all intensive purposes, the term "lottery" shall be used only in reference to the UK national lottery. This paper intentionally differentiates between household and personal income as measures of income in its analysis to account for the different measures of income used in different literatures. A measure of general lottery participation in the past year is considered as the main binary dependent variable for analysis before a more qualitative binary dependent variable of **monthly** lottery participation is chosen instead to reveal *consistent* risk seeking behavior. A probit model will be the main econometric model considered. This paper argues that while household income is a superior measure of general wealth, the inclusion of personal income variables could potentially reveal measures of income satisfaction that avoids issues of self reported bias.

Literature Review

Theoretical Background

According to classical expected utility theory, the act of gambling is often irrational as individuals willingly partake in activities that yield negative expected returns. Yet lottery gambling activity has always been prevalent in human history. To explain this discrepancy, many theories in behavioral economics have been developed, particularly in relation with one's income and wealth.

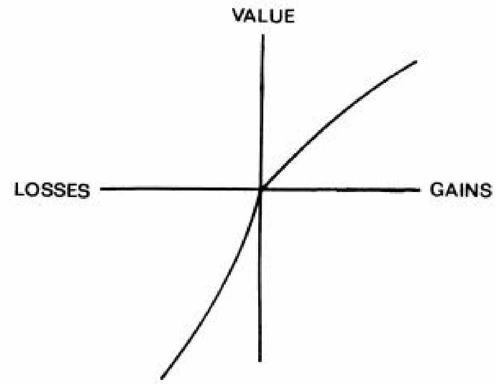
Friedman Savage Model



Graph by Brunk (1981)

Friedman and Savage (1948) describes why individuals often contradictorily partake in both risk averse (buy insurance) and risk seeking (buy lottery tickets) behavior simultaneously by developing a utility function that gradually changes from convex to concave to convex again over increasing levels of *absolute* wealth. Their hypothesis particularly emphasizes on the gambler's desire to reach higher socio-economic status through the middle transitory concave section of the function, where gamblers are more risk seeking. The curvature of their model suggests those with middle levels of income tend to engage most in gambles that incur small costs and high rewards with low probability of winning, where costs are small enough that it does not reduce wealth to that before the first convex segment, thus implying an "upside down U shape" effect of income on lottery participation.

Prospect Theory



Graph by Kahneman, Tversky (1979)

Tversky and Kahneman (1979) on the other hand modeled the effects of relative income on gambling behavior with prospect theory. They derived an S-shaped value function which is concave in gains and convex in losses with a general aversity to loss. The shape of the function suggests how individuals are more risk seeking in losses and risk averse in gains in reference to a net zero point of income. It is important to note that the value function aligns with agents' perceived beliefs of outcomes instead of actual probabilities, and are hence subject to misinterpretation (Thaler 2015). This explains the agent's seemingly irrational behavior of partaking in negative expected gambling due to an overvalued probabilistic belief of winning. However, a key flaw with this model is how it neglects psychological aspects of gambling (Shafir and LeBoeuf 2002), such as how perceived feelings of gain and loss prior to the gamble (external effects like income and social status) may cause an individual to begin at a starting point that is not equal to zero, but instead a point along the S shaped curve. These psychological factors must be accounted for when considering the causal effects of income in lottery participation.

Empirical Literature

Brunk's (1981) empirical test in the US on the Friedman and Savage model suggest the uniqueness of lottery gambling compared to other forms of gambling. He argues how lottery is one of the few forms of gambling that offers the low risk high reward payoff that the Friedman-Savage model suggests as a motivation to move up the socioeconomic ladder. His regression on gambling participation uniquely included a self-reported "income dissatisfaction" variable, and found how those who are more dissatisfied with their income tend to significantly participate more in lotteries, implying the validity of the Friedman-Savage model. This variable becomes insignificant when considering other forms of gambling(bingo, raffles poker), highlighting the social motivations of gambling in those activities that is less apparent in lotteries. However, the sample size of this research is relatively small (n=147), and may not be indicative of gambling behavior of the mass majority. The measure of “income dissatisfaction” may also be subjected to issues of unreliable self reporting, causing inaccurate interpretations to be made.

Most empirical literature investigate effects of income on lottery gambling behavior under two different terms: household income and personal income. Interestingly, both income effects tend to be studied individually, studies opting to research effects of either one as a general term of “income” rather than controlling for both simultaneously. Both income variables are similar in nature, with personal income measuring how much an individual earns and household income measuring how much an entire household earns over a period of time.

On household income, Beckert and Lutter (2008) researched the effects of household income on lottery gambling in Germany. A probit model on a dataset that includes both gamblers and non-gamblers found how household income is statistically insignificant in affecting lottery participation, yet further restricted OLS regressions on lottery expenditure from gamblers found a significant, positive relationship with household income, where those with higher income tends to have higher *absolute* expenditure, but lower *proportional* expenditure on lottery gambling. Moreover, further demand analysis found how gamblers in the lower middle classes with mostly average absolute incomes tend to play the most, which matches with the concave segment in the Friedman and Savage model that predicts how lower middle income groups tend to be the most risk-seeking.

On personal income, Orford et al. (2010) researched the effects of personal income on gambling frequency using data from the 2007 UK gambling prevalence survey. Their results found that personal income was positively and linearly related to gambling frequency and participation. Note this paper draws no distinction between types of gambling activities in its analysis, which is problematic as different gambling activities may insinuate different gambling behavior in relation to income (the social distinction between lotteries and other gambling activities). However, this paper still suggests that gambling is a luxury activity that increases as personal income increases, though whether that translates the same for lotteries remain to be investigated.

Other controls of gender and employment status are found to be significant across most research on gambling[2], which must be controlled for in my analysis.

[2] Males and those employed are found to gamble more

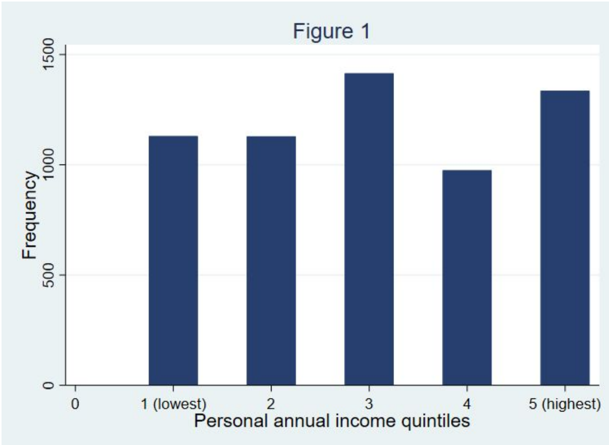
Data

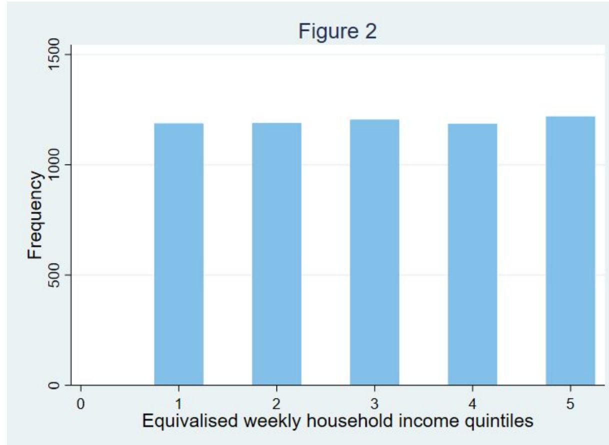
This paper uses data from the British Gambling Prevalence Survey (2010) that is conducted by the National Centre of Social Research, which includes data on participation in all forms of gambling in the UK.

This dataset was chosen based on two criterion: sample size and assessment method. This survey was more targeted towards the study of gambling behavior and includes a larger sample size of gamblers for analysis., Hence, it is chosen over more recent health surveys in 2012 and 2016. In terms of assessment method, this survey was conducted using computer assisted self interviewing, minimizing problems of social desirability bias that may cause individuals to systematically underreport gambling behavior, as there is often negative stigma towards addictive gambling. Negative terms such as “gambling” were also avoided in the questionnaire to avoid negative biases and replaced with terms like “Leisure Time” and “Lottery”.

The dataset consists of 5,985 observations after irrelevant or faulty data was removed[3]. Both income variables are divided into five quintiles, where trends comparing the two variables of income shall be plotted across their respective dummy variables. Personal income and household income are also measured annually and weekly respectively. A direct timeframe comparison would have been preferred where both income variables are both measured annually or weekly, but is unavailable. This deficiency is mitigated given that one’s weekly household income is unlikely to change significantly throughout a given year. However, this is still an assumption. It is important to note that the household income measure in the dataset uses *equivalised* household income, bands of household income that is weighted on household size and composition under the widely used McClements scoring system to make sensible income comparisons between households (Anyaegbu 2010).

[3] Such as those who have refused to report income figures





Figures 1 and 2 indicate the distribution of personal and household income quintiles in the dataset. Household income is relatively evenly distributed, while personal income quintiles are somewhat skewed towards the third and fifth quintiles, though it is still sufficiently evenly distributed to avoid selection bias.

Table 1

Household/Personal	1 (lowest)	2	3	4	5 (highest)
1 (lowest)	476	553	117	25	16
2	246	218	459	209	57
3	196	138	472	211	187
4	181	183	270	260	292
5 (highest)	31	37	97	270	784

Table 1 illustrates the frequency distribution of combinations of household and personal income quintiles of individuals. The diagonal has been highlighted to show that the most frequent combinations of income tend to be those with equal or similar household/personal income quintiles. However, while related, the relationship between household and personal income quintiles is not perfectly collinear, with 22.3% of surveyed individuals outside the yellow and green diagonal, suggesting limited multicollinearity.

Table 2[4]

Variable	Mean	s.d.	Min/Max
Lottery Participation	0.619	0.486	0-1
Monthly Lottery Participation	0.483	0.500	0-1

Summary statistics of lottery participation in table 2 indicate that 61.9% of individuals have bought national lottery tickets in the past year. While general lottery participation could be a suitable dependent variable used for this study, an ideally more qualitative measure of lottery participation that can reveal tendencies for *consistent* risk seeking behavior is preferred. An example for a more qualitative measure of gambling participation is gambling frequency. However, although initially considered, measures of exact gambling frequencies are problematic, as individuals are unlikely to be able to accurately report the exact number of days/hours spent gambling in the past year, and hence are more heavily influenced by social desirability bias in self reporting. Therefore, a binary dependent variable for monthly lottery participation is chosen to be the main focus of this study instead, as it includes a qualitative aspect to the dependent variable while also general enough to minimize bias in self-reporting. 48.3% of the dataset reported to have participated monthly in lottery gambling, implying a sufficiently sized sample for analysis.

[4] Full summary statistics in appendix II shows a relatively even distribution of those employed/unemployed and gender

Methodology:

As the dependent variable is binary, an OLS Linear Probability Model(LPM) is initially considered. However, a LPM is subject to flaws of non-normality, heteroskedasticity and value ranges outside the 0-1 bound (Greene 1996). A probit/logit regression gives more consistent estimates for binary dependent variables (Cameron and Trivedi 2005). While mostly similar, a probit model is chosen over logit for ease of comparison with Beckert and Lutter's paper.

The general probit model specification is as follows:

$$\Pr(m\_lot_i = 1) = \phi[\alpha + \beta_1 pincome_i + \beta_2 hincome_i + \sum_{k=3}^{K} \beta_k controls_{ik} + \varepsilon_i]$$

Where *m_lot_i* refers to a dummy variable that equals 1 if an individual has participated monthly in lottery gambling, 0 otherwise. Variables *$pincome_i$* refers to the second to fifth personal income quintile dummies, while *$hincome_i$* refers to the second to fifth household income quintile dummies that will be used in the regression. The first quintiles (lowest) of each income variable is omitted to avoid perfect collinearity. A full list of controls can be found in appendix I.

Two additional regressions are run in conjunction with the general model, one which does not control for household income, while the other does not control for personal income.

<u>Multicollinearity</u>

A key concern with this econometric design is the potentially collinear relationship between household and personal income, where richer individuals tend to be in richer households. This would increase standard errors and cause estimators to be less precise.

While table 1 has shown that the relationship between household and personal income is not perfectly collinear, a Variance Inflation Factor (VIF) Test will be used on the LPM model of m_lot with the same independent and controlled variables. Should the problem of multicollinearity prove to be insignificant in the OLS model, it is possible to infer the same conclusion for the probit model.

<u>Omitted Variables</u>

Three fundamental relevant variables are unreported in the dataset, and hence omitted from the regression.

Windfall Income:

Papers by Buddelmeyer et al. (2014) and Christiansen (1998) found that windfall income - an unexpected, transitory increase in income from sources such as fiscal stimulus - increases general gambling participation. This may bias results where middle-lower income groups are shown to gamble significantly more due to a sudden exogenous increase in *proportional* income. However, this is mitigated given how gains in windfall income are (arguably) random, and hence may not bias estimates.

Risk Attitudes:

Pre-existing or hereditary effects such as a more easily activated limbic system may cause individuals to respond preferentially to immediate rewards than future rewards, leading to higher likelihood of addictive gambling behavior (McClure et al. 2004). Controls of dummies for smoking status and alcohol consumption are included to mitigate this problem, as they are commonly accepted proxies that can reveal preexisting risk attitudes. Moreover, studies such as one by Cook et al. (2015) describe a causal relationship between substance abuse and tendency for addictive behavior, suggesting how alcohol and smoking dummies can act as efficient proxies.

Peer Effects

Individuals with a higher number of gambling peers may gamble more. This problem, while partially mitigated by the inclusion of parental gambling behavior, is not fully resolved as it neglects gambling behavior of other peers such as those of friends and spouses. However, as people participate in lottery gambling less so for social reasons, omitted peer effects are less concerning.

Results and Interpretation

Variable	Monthly Lottery Participation Probit Marginal Effects		
	(1)	(2)	(3)
Household Income			
Quintile 2		0.0362 (0.121)	0.0107 (0.663)
Quintile 3		0.0223 (0.344)	-0.00748 (0.765)
Quintile 4		-0.0310 (0.197)	**-0.0633**** (0.014)
Quintile 5		**-0.0767***** (0.002)	**-0.130***** (0.000)
Personal Income			
Quintile 2	**0.0473**** (0.045)		**0.0437*** (0.065)
Quintile 3	**0.0841***** (0.001)		**0.0887***** (0.000)
Quintile 4	**0.0782***** (0.006)		**0.112***** (0.000)
Quintile 5	**0.0446*** (0.093)		**0.114***** (0.000)

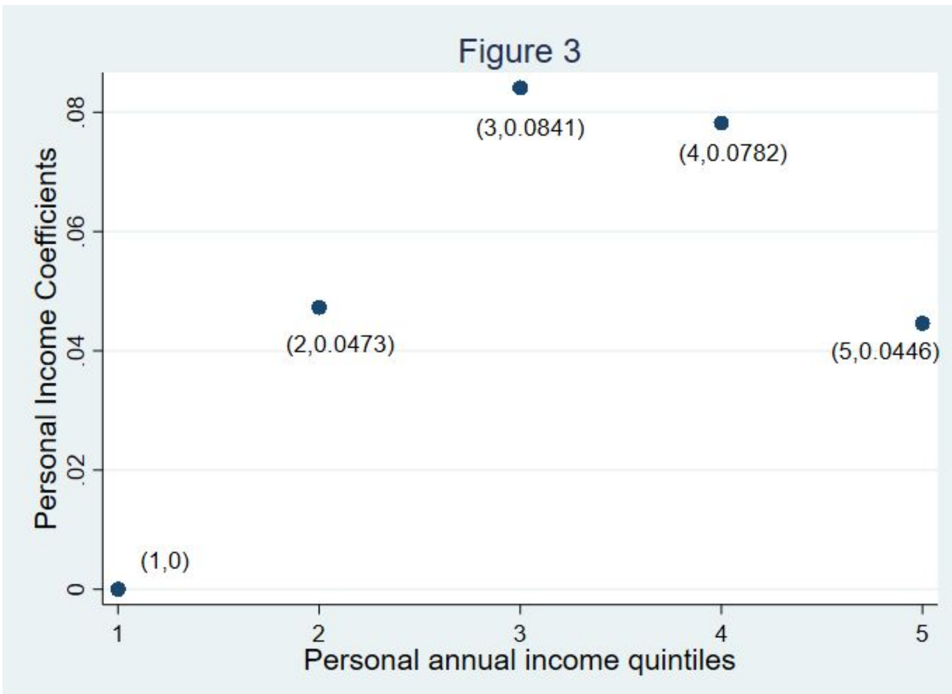
* p < 0.10, ** p < 0.05, *** p < 0.01
P values in parentheses
N = 5985

The above regression table reports the probit marginal effects of personal and household income[5]. Two regressions that include only household or personal income quintiles are considered before a regression that includes both income variables is conducted.

[5] Controls are included in all 3 regressions but not reported. Full regression table in Appendix VI & VII

Regression 1

Regression 1 illustrates the marginal effects of personal income on tendencies for monthly lottery participation in the past year **without** holding household income constant. Interpretation of the second personal income quintile is how individuals in the second personal income quintile are on average 4.73% more likely to have participated in the lottery monthly than individuals in the first personal income quintile, ceteris paribus. The remaining marginal personal income coefficients can be interpreted similarly.



Results in regression 1 suggest how individuals in the second to fifth personal income quintile are *significantly* more likely to have participated in lotteries monthly in the past year than those in the first quintiles, implying a positive, upside down U-shape relationship of personal income with lottery participation as seen in figure 3. This matches closely with the Friedman-Savage model, which argues how individuals with middle levels of income tend to participate more in gambles that would lead to higher levels of socio-economic status through a middle transitory concave section of the function. A stronger economic motive to transition to a higher socioeconomic class through lotteries may explain why middle class individuals choose to participate in risk seeking behavior more frequently than those in the highest personal income quintiles who have already achieved said status. On the other hand, greater financial ability to bear risks of losing may also explain why they gamble more than those with lower income, thus accounting for why personal income's effects are skewed to the middle.

These findings slightly differ with findings of Orford et al. (2010) from the 2007 UK gambling prevalence survey, where a linearly positive relationship between personal income and general gambling frequency is observed. This may be due to the aforementioned lack of specification in gambling activities in their paper. For example, social gambling may be more prevalent in individuals with higher personal income, while more economically motivated gambles (e.g. lotteries) tend to be less prevalent among higher earning individuals. This lack of distinction may be a source of the discrepancy in results.

However, this trend may also be observed due to omitted variables like windfall income, which as mentioned could bias results that cause middle income groups to gamble more due to a higher gain in proportional income. Furthermore, the discrepancy in results with Orford et al. (2010) may be due to failure to control for 'household social position' (controlled for in household income variables), which is found to be statistically significant in their paper in affecting gambling participation. This is especially relevant when considering how a Ramsey RESET test on an OLS LPM equivalent of regression 1 rejects the hypothesis of no omitted relevant variables at the 5% significant level[6]. While OLS diagnostic results are not necessarily perfectly translatable into a probit model, it is still possible to infer a possibility of omitted relevant variables in regression 1.

<u>Regression 2</u>

Regression 2 illustrates the marginal effects of household income on tendencies for monthly lottery participation in the past year **without** holding personal income constant. Interpretation of the second household income quintile is how individuals in the second household income quintile are on average 3.62% more likely to have participated in the lottery monthly than individuals in the first household income quintile, ceteris paribus. The remaining marginal household income coefficients can be interpreted similarly.

Results in regression 2 suggest that household income plays a mostly insignificant role in affecting tendencies for regular lottery participation, where only individuals in the highest household income quintile participate significantly less than other individuals. This implies those in the poorest four household income groups have the same tendencies to participate regularly in lotteries. The results somewhat matches Beckert and Lutter's (2008) findings in Germany of household income's general insignificance in overall participation. The strong negative significance of the fifth household income quintile could also explain the trend observed in their paper of how gamblers with

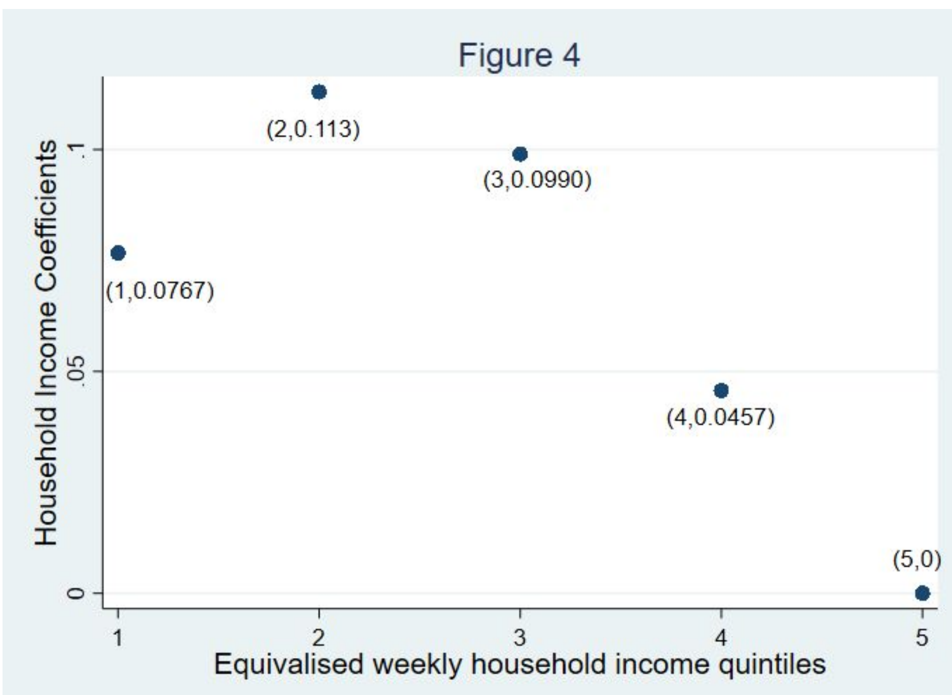
[6] Appendix III

middle to lower household income participate in lotteries more frequently than those with higher incomes, although unlike Beckert and Lutter's regression, regression 2 is not restricted to consider only gamblers.

Regardless, the similarities with Beckert and Lutter's (2008) results also implies how the household income variables somewhat reflect the effect of wealth described in the Friedman Savage model, though unlike the personal income variables in regression 1, the first income quintile is included as part of the "middle transitory concave function". This observation is further validated when the fifth household income quintile (instead of the first) is omitted and a strongly significant reverse U-shape trend can be observed as seen in figure 4, similar to personal income.

Household Income Quintiles	1	2	3	4
Marginal Effects	**0.0767*** ** (0.002)	**0.113*** ** (0.000)	**0.0990*** ** (0.000)	**0.0457** ** (0.038)

* p < 0.10, ** p < 0.05, *** p < 0.01
P values in parentheses



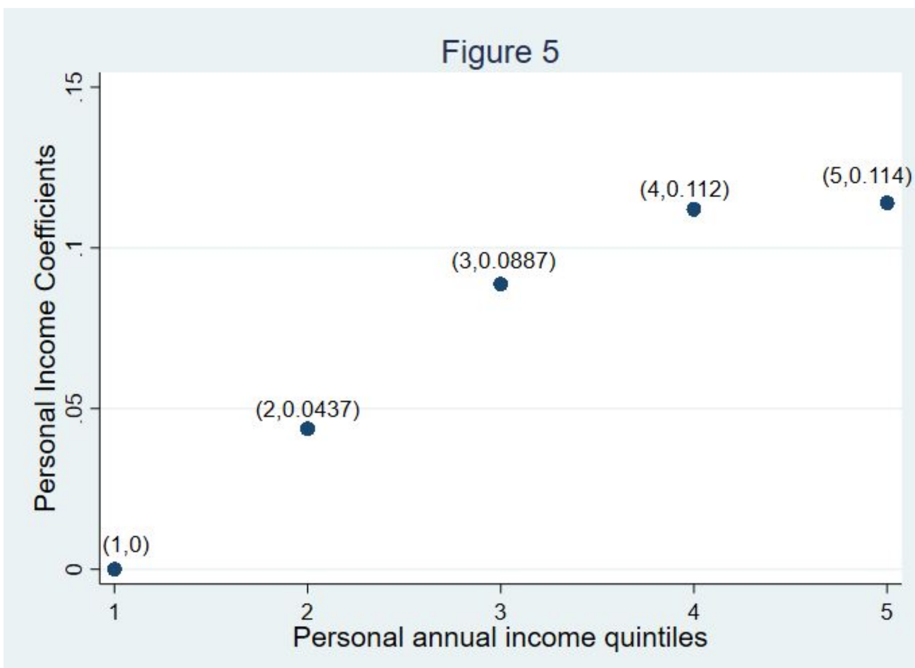
The inclusion of the first household income quintile as part of the middle transitory risk seeking segment of the Friedman Savage model for lottery participation makes sense as lottery gambling bears little risk that makes it unlikely for individuals to fall into the first risk averse, convex segment from excessive lottery gambling. Interestingly, a

Ramsey RESET test on the LPM equivalent of regression 2 failed to reject the hypothesis of no omitted relevant variable at the 5% significance level[7], suggesting how household income is a better independent variable than personal income to measure tendencies for monthly lottery participation in the past year.

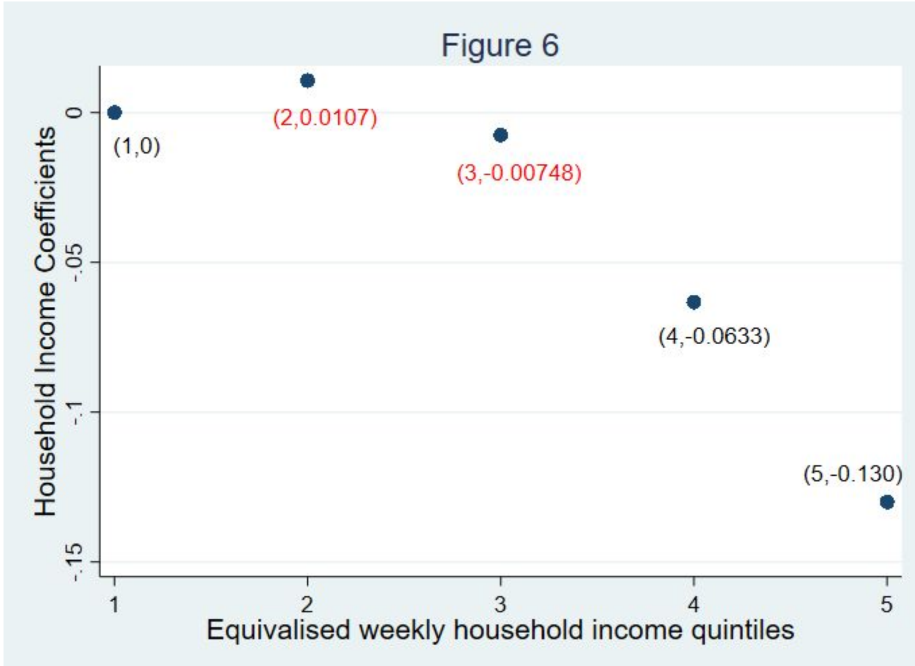
Regression 3

Regression 3 illustrates the marginal effects of both personal and household income on tendencies for monthly lottery participation in the past year while holding each income variable constant. Personal income coefficients describe average differences in likelihood for monthly lottery participation in the past year with those in the first personal income quintile while holding household income constant, ceteris paribus. Household income coefficients describe average differences in likelihood for monthly lottery participation in the past year with those in the first household income quintile while holding personal income constant, ceteris paribus.

Suggestive diagnostic tests run on the LPM equivalent of regression 3 show no strong evidence for omitted relevant variables[8]. A VIF test[9] also found no strong evidence for multicollinearity between both income variables and other controls.



[7] Appendix III
[8] Appendix III
[9] Appendix IV



*Red label implies insignificant income quintile

Once household income is controlled for, the trend in personal income matches much closer with trends found by Orford et al. (2010): a positive, linear relationship that increases with gambling frequency. This implies that lottery gambling, similar with other forms of gambling is a luxury activity that increases as personal income increases. However, this trend is counterintuitive in a few ways.

As mentioned, unlike other forms of social gambling, lotteries are mostly economically motivated. The current personal income trend implies those who are richer have a stronger economic incentive than the poor to win the lottery, but this does not make sense when assuming a diminishing marginal utility of wealth that is captured in the second convex segment in the Friedman Savage model. This makes even less sense when considering how the trend in household income in the same regression is increasingly negative, suggesting seemingly opposite income effects.

To understand this, it is important to remember that the linear trend of personal income is observed only when holding household income constant. Hence, the trend observed implies that tendencies to participate in lotteries monthly only increases linearly as **relative wealth to household increases**. This distinction is important as the linear trend observed in personal income in regression 3 can now be interpreted as a similar measure of income satisfaction that is used by Brunk (1981), who found those who are more dissatisfied with current level of household income tend to participate more in lotteries.

The positive, linear trend in personal income trend makes sense under this context. Individuals who earn more personally may have a higher expectation to reach a higher socio-economic status. Yet when household income, which is a better measure of real spending power as it accounts for spending responsibilities[10], falls below said expectation, individuals have higher motivations to reach higher status by playing lotteries, especially when discrepancy between personal income and real spending power increases. This measure could potentially be more accurate than Brunk's (1981) variable of income satisfaction, as personal income variables avoid problems of subjectivity in self reporting while also revealing measures of income satisfaction, further justifying his claim.

The reverse U-shaped trend in household income with a peak in the second quintile remains in regression 3 despite additional controls of personal income, again matching predictions of income effects in the Friedman Savage model and suggesting household income's robustness as an indicator of wealth and spending power. A prospect theory model that incorporates psychological factors could also potentially explain household income's mostly negative relationship with gambling frequency. As mentioned, while standard prospect theory weighs risk attitudes at a starting position of zero, external effects like income and social status may affect perceived feelings of gain and loss prior to the gamble. For example, individuals with lower household income may suffer from higher perceived feelings of loss and gamble at a starting point to the left on the convex segment of the S-shaped curve, hence adopting more risk seeking behavior. However, as household income increases, perceived feelings of gains increases, causing individuals to gamble at a starting position to the right on the concave segment of the S-shaped function and adopting more risk averse behavior, thus accounting for the increasing aversity to risk as household income increases and accounting for its negative trend. The effects of subjective relative income as a motivation for lottery participation has also been suggested by Haisley et al. (2008).

Post Estimation Robustness Checks

On top of suggestive diagnostic tests on LPM equivalents for regressions 1 to 3, ROC curves were also used to test the predictive power of all regressions[11]. Results show mostly average but similar explanatory strengths of each regression. Predictive power

[10] For example, individuals who earn high personal income but is the sole breadwinner of a large household may not be able to spend as much, whereas individuals who earn low personal income but is in a wealthy household may spend more

[11] Appendix V

may be improved should the aforementioned unavailable omitted relevant variables are included as controls.

Conclusion and Limitations

Findings illustrate effects of income on regular lottery participation. In a regression that controls for both household and personal income, personal income potentially illustrates a revealed measure of income dissatisfaction that increases as personal income creases while holding household income constant. Household income, other other hand, matches with the trend hypothesized in the Friedman Savage model that illustrates a reverse U-shaped relationship between household income and tendencies for regular lottery participation even while holding personal income constant. Should only one income variable be considered when testing the relationship between lottery participation and income (such as in the literature reviewed), this paper suggests household income to be more indicative as a variable of wealth due to its consistency in trends across both included regressions and a failure to reject the hypothesis that there are no omitted relevant variables in the regression that only includes household income as the income variable, both contrary to findings for personal income.

However, there are limitations to my study. Other than the previously mentioned problems of omitted variables, exact, absolute values of household and personal income would make the observed trend more convincing as compared to the five data points available for each income quintile. A square term to account for the maximum middle income coefficient could be used to test whether a reverse U-shape could still be observed under this model.

Given accurate self reporting, a more exact measure of gambling frequency (such as number of days gambled in the past year) as a dependent variable would also make findings more accurate. A Heckman Two-Step model could be adopted to avoid the problem of left censored data by building upon an initial probit model of lottery participation before weighing it with the same effect as an omitted variable on an OLS regression for lottery participation frequency, though this method is beyond the scope of this study[12].

Admittedly, the interpretation of the personal income variable in regression 3 as a variable of income satisfaction is somewhat speculative. Ideally, a more accurate measure of income satisfaction that avoids problems of self-reporting bias would be

[12] A tobit model would also work for censored data, but fails here as tobit models are best suited for continuous dependent variables rather than discrete measures of participation frequency (Amore and Murtinu 2019)

better preferred to test whether Brunk's (1981) findings matches with results in this study. As Haisley et al. (2008) also mentioned the importance of relative income to neighborhood and peers as important indicators of income satisfaction over absolute personal income, the argument for relative income to household as a factor of income satisfaction is only sound when the same trend and significance of personal income can be observed even when additional controls of relative income is included (which is unavailable in this dataset).

Policy Implications and Extensions

The results on this study implies the strength of the Friedman Savage model in predicting behavior for low risk high reward gambles in relation to one's wealth. The fact that low (household) income individuals gamble more frequently than those with high incomes suggest how government operated lotteries are by nature a form of regressive tax, assuming that higher participation rate in lotteries imply higher spend, which was implied in findings by Castrén et al. (2018). Should this hold true, then revenue gained from the national lottery should be specifically allocated to aid those of lower incomes to combat the regressive nature of lotteries and improve equity.

As extensions to this study, it would be interesting to observe whether the same trend could be observed in gambles that involve higher risk high reward games contrary to the low risk high reward structure of lotteries. Under this scenario, the Friedman Savage model predicts that middle income individuals would participate most in these games, as low income individuals are unable to bear the risk involved and remains in the first convex segment of the function. International comparisons of lottery participation, such as gambling prevalence in Macau's Mark Six Lottery (Fong 2005) would also further attest to the universal strength and validity of the Friedman Savage model.

Appendix

Appendix I: Variable List

Variable	Description
Personal Income Quintiles (1st quintile lowest, 5th quintile highest)	Personal Income Quintile Dummies 2 to 5. Equals 1 if respondent is in specified quantile, 0 otherwise: - 5th Quintile - 4th Quintile - 3rd Quintile - 2nd Quintile (1st Quintile omitted)
Household Income Quintiles (1st quintile lowest, 5th quintile highest)	Household Income Quintile Dummies 2 to 5. Equals 1 if respondent is in specified quantile, 0 otherwise: - 5th Quintile - 4th Quintile - 3rd Quintile - 2nd Quintile (1st Quintile omitted)
Age	Age (years) of respondent at time of survey
Age^2	Square term of age
Gender	Equals 1 if individual is male, 0 otherwise
Education	Highest educational qualification dummies. Equals 1 if respondent matches specified qualification, 0 otherwise: - GCSE - A-level - Professional Degree - University Degree - Others (No qualification omitted)
Smoking Status	Equals 1 if individual smokes, 0 otherwise
Drinking Status	Equals 1 if individual drinks alcohol, 0 otherwise
Ethnicity	Ethnicity dummies. Equals 1 if respondent is in specified ethnic group, 0 otherwise - Mixed - Asian/Asian British - Black/Black British - Chinese/Other

	(White omitted)
Parental Gambling Behavior	Parental gambling behavior dummies. Equals 1 if respondent's parents matches specified behavior, 0 otherwise - Parents are problem gamblers - Parents are regular gamblers, but not problem gamblers (Parents did not gamble omitted)
Region	Region dummies. Equals 1 if respondent lives in specified region, 0 otherwise - Wales - Scotland (England omitted)
Marital Status	Equals 1 if respondent is married, 0 otherwise
Economic Status	Economic status dummies. Equals 1 if respondent matches specified economic group, 0 otherwise - Retired - Student - Unemployed - Hdo (health disability, home caretaker, other) (Employed omitted)
Ads	Equals 1 if respondent recalls seeing advertisements for gambling, 0 otherwise
Household Size	Number of individuals living in respondents' household
Household Economic Status	Equals 1 if respondent's household generates income (in addition to respondent), 0 otherwise

Appendix II: Summary Statistics

Variable		Mean	Standard Deviation	Min/Max
Personal Income	Quintile 1	0.189	0.391	0-1
	Quintile 2	0.189	0.391	0-1
	Quintile 3	0.236	0.425	0-1
	Quintile 4	0.163	0.369	0-1
	Quintile 5	0.223	0.416	0-1
Household Income	Quintile 1	0.198	0.399	0-1
	Quintile 2	0.199	0.399	0-1
	Quintile 3	0.201	0.401	0-1
	Quintile 4	0.198	0.397	0-1
	Quintile 5	0.204	0.403	0-1
Age		46.830	17.831	16-97
Gender		0.463	0.499	0-1
Education	GCSE	0.240	0.427	0-1
	A-levels	0.121	0.326	0-1
	Professional Degree	0.0744	0.262	0-1
	University Degree	0.317	0.466	0-1
	Others	0.0165	0.128	0-1
	No Qualification	0.23	0.421	0-1
Smoking Status		0.250	0.433	0-1
Drinking Status		0.629	0.483	0-1
Ethnicity	Mixed	0.0102	0.100	0-1

	Asian/Asian British	0.0391	0.194	0-1
	Black/Black British	0.0246	0.155	0-1
	Chinese/Other	0.00735	0.0854	0-1
	White	0.919	0.273	0-1
Parental Gambling Behavior	Problem Gamblers	0.0389	0.193	0-1
	Regular Gamblers	0.209	0.407	0-1
	Non-gamblers	0.752	0.432	0-1
Economic Status	Retired	0.206	0.405	0-1
	Student	0.0642	0.245	0-1
	Unemployed	0.0294	0.169	0-1
	hdo	0.148	0.355	0-1
	Employed	0.553	0.497	0-1
Lottery Participation		0.619	0.486	0-1
Monthly Lottery Participation		0.483	0.500	0-1

Appendix III: Omitted Relevant Variable Test

Ramsey RESET test run on OLS LPM equivalents of each regression to test for omitted relevant variables:

H0: model has no omitted variables

Regression	**F statistic**	**Prob>F**	**Reject null at 5% significance level?**
General Model			
1	F(3,5951) = 2.65	0.0473	Reject

2	F(3,5951) = 1.48	0.2183	Do not reject
3	F(3,5947) = 2.07	0.1003	Do not reject

Conclusion: Results suggest only model 1 (considers only personal income) have problems of omitted relevant variable that may translate across OLS and probit models.

Appendix IV: Multicollinearity Test

Variance Inflation Factor(VIF) Multicollinearity Test run on OLS LPM equivalents of each regression 3 test for multicollinearity

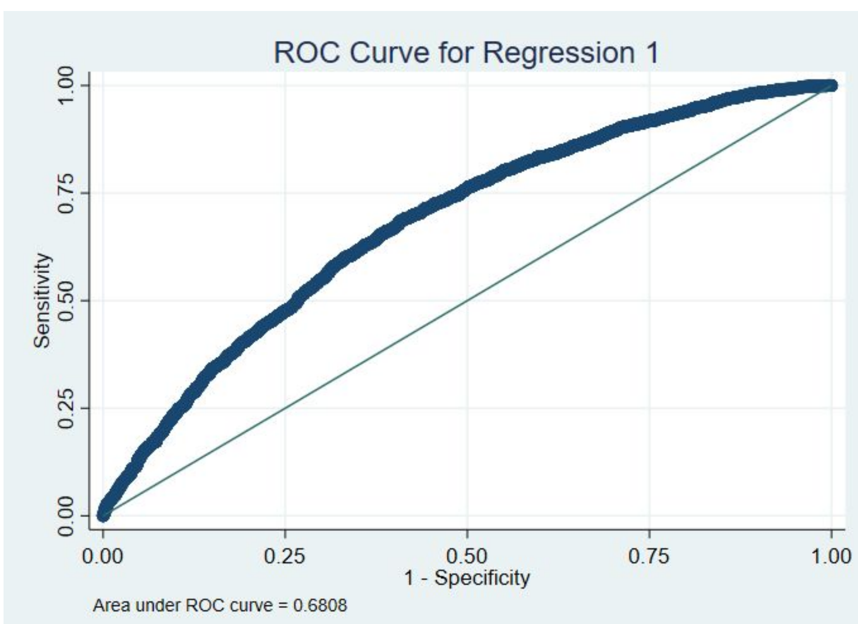
Model	Mean VIF Value
3	1.83

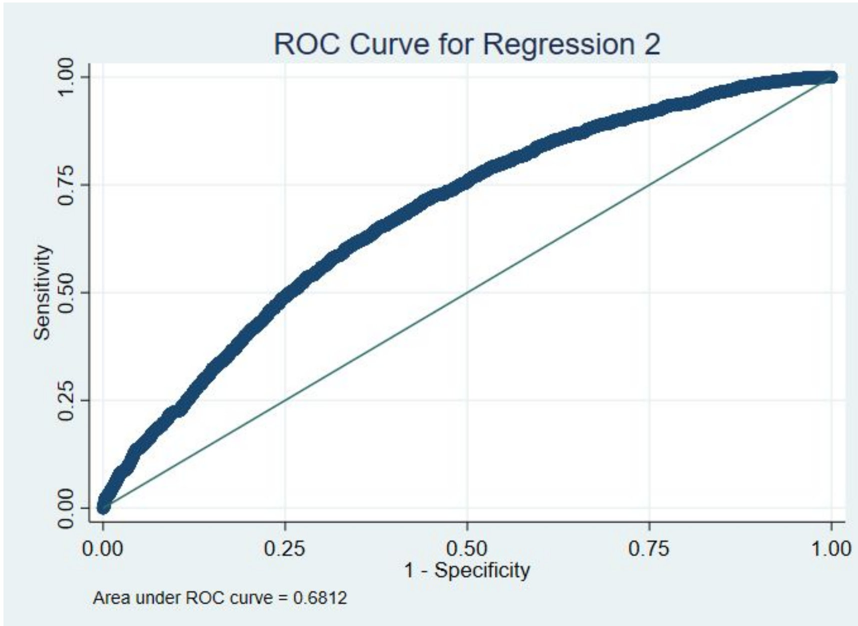
Age^2 term removed in VIF test as its collinear relationship with age term is expected and artificially increases VIF

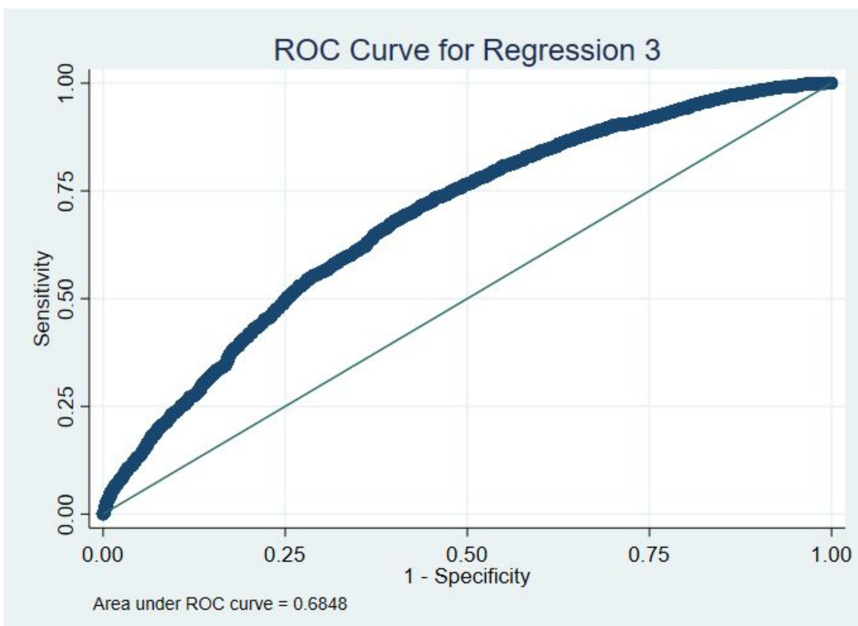
Conclusion: As mean VIF<10, no strong evidence of multicollinearity can be inferred from OLS LPM diagnostics for regression 3

Appendix V: Postestimation Robustness Check

Predictive Power: An area of 1 indicates perfect predictive power of the model. Area of 0.5 suggests no predictive power at all. Results suggest similar and average predictive power across all 3 models.




ROC Curve for Regression 2
Sensitivity
0.00
0.25
0.50
0.75
1.00
1 - Specificity
Area under ROC curve = 0.6812


ROC Curve for Regression 3
Sensitivity
0.00
0.25
0.50
0.75
1.00
1 - Specificity
Area under ROC curve = 0.6848

Appendix VI: Full Probit Regression Results (Coefficients)

Number of observations =5985

Variable	1	2	3
Household Income			
Quintile 2		0.0909 (0.0586)	0.0268 (0.0616)
Quintile 3		0.0559 (0.0591)	-0.0188 (0.0629)
Quintile 4		-0.0779 (0.0603)	**-0.159**** (0.0644)
Quintile 5		**-0.193***** (0.0632)	**-0.326***** (0.0753)
Personal Income			
Quintile 2	**0.119**** (0.0593)		**0.110*** (0.0594)
Quintile 3	**0.211***** (0.610)		**0.223***** (0.0631)
Quintile 4	**0.196***** (0.688)		**0.280***** (0.0720)
Quintile 5	**0.112*** (0.0666)		**0.286***** (0.0754)
age	**0.0531***** (0.00614)	**0.0540***** (0.00611)	**0.0513***** (0.00617)
age2	**-0.000491***** (0.0000638)	**-0.000501***** (0.0000638)	**-0.000481***** (0.0000641)
gender	**0.128***** (0.0360)	**0.141***** (0.0350)	**0.104***** (0.0363)
degree	**-0.575***** (0.0555)	**-0.509***** (0.0548)	**-0.552***** (0.0560)
professional	**-0.316***** (0.0733)	**-0.266***** (0.0729)	**-0.308***** (0.0736)
A Level	**-0.257***** (0.0675)	**-0.221***** (0.0674)	**-0.247***** (0.0678)
GCSE	**-0.129**** (0.0543)	**-0.102*** (0.0543)	**-0.119**** (0.0545)
others	-0.198 (0.408)	-0.183 (0.134)	-0.200 (0.134)
smoke_dummy	**0.193***** (0.0408)	**0.183***** (0.0410)	**0.183***** (0.0410)
mixed_eth	0.0102 (0.167)	-0.0181 (0.166)	-0.0150 (0.167)
asian_ab	**-0.475*****	**-0.476*****	**-0.479*****

	(0.0975)	(0.0978)	(0.0980)
black_bb	**-0.203*** (0.113)	**-0.215*** (0.112)	**-0.216*** (0.112)
chinese_other	**-0.502**** (0.221)	**-0.522**** (0.223)	**-0.512**** (0.222)
prgamble	**0.229***** (0.0418)	**0.231***** (0.0418)	**0.231***** (0.0418)
ppgamble	**0.288***** (0.0908)	**0.281***** (0.0910)	**0.281***** (0.0910)
drinker	0.0225 ((0.0378)	0.0359 (0.0379)	0.0349 (0.0380)
married	0.0640 (0.0427)	0.0697 (0.0437)	**0.0955**** (0.0444)
Wales	-0.0399 (0.0798)	-0.0502 (0.0798)	-0.0543 (0.0801)
Scotland	0.0370 (0.0554)	0.0325 (0.0555)	0.0307 (0.0556)
retired	**-0.153*** (0.0828)	**-0.207**** (0.0819)	-0.129 (0.0839)
hdo	**-0.178***** (0.0688)	**-0.265***** (0.0645)	**-0.162**** (0.0690)
student	**-0.456***** (0.101)	**-0.585***** (0.0952)	**-0.437***** (0.101)
unemployed	0.0899 (0.113)	-0.0503 (0.108)	0.0870 (0.113)
ads	**0.0895*** (0.0495)	**0.0945*** (0.0495)	**0.0904*** (0.0496)
hsize	-0.0253 (0.0160)	**-0.0615***** (0.0174)	**-0.0690***** (0.0178)
hemployed	0.0574 (0.0619)	0.0483 (0.0630)	0.0854 (0.0637)
constant	**-1.325***** (0.176)	**-1.105***** (0.174)	**-1.162***** (0.180)

* p < 0.10, ** p < 0.05, *** p < 0.01
Robust standard errors in parentheses

Appendix VII: Full Probit Regression Results (Marginal Effects)

Number of observations =5985

Variable	1	2	3
Household Income			
Quintile 2		0.0362 (0.0233)	0.0107 (0.0245)
Quintile 3		0.0228 (0.0235)	-0.00748 (0.0250)
Quintile 4		-0.0310 (0.0240)	**-0.0633**** (0.0257)
Quintile 5		**-0.0767***** (0.0252)	**-0.130***** (0.0300)
Personal Income			
Quintile 2	**0.0473**** (0.0236)		**0.0437*** (0.0236)
Quintile 3	**0.0841***** (0.0234)		**0.0887***** (0.0251)
Quintile 4	**0.0782***** (0.0274)		**0.112***** (0.0287)
Quintile 5	**0.0446*** (0.0265)		**0.114***** (0.0300)
age	**0.0211***** (0.00245)	**0.0215***** (0.00243)	**0.0204***** (0.00246)
age2	**-0.000196***** (0.0000254)	**-0.000200***** (0.0000254)	**-0.000191***** (0.0000255)
gender	**0.0510***** (0.0143)	**0.0561***** (0.0139)	**0.0415***** (0.0145)
degree	**-0.229***** (0.0221)	**-0.203***** (0.0218)	**-0.220***** (0.0223)
professional	**-0.126***** (0.0292)	**-0.106***** (0.0290)	**-0.123***** (0.0293)
A Level	**-0.102***** (0.0269)	**-0.0880***** (0.0269)	**-0.0986***** (0.0270)
GCSE	**-0.0513**** (0.0216)	**-0.0405*** (0.0216)	**-0.0472**** (0.0217)
others	-0.0788 (0.0532)	-0.0728 (0.0533)	-0.0799 (0.0532)
smoke_dummy	**0.0767***** (0.0163)	**0.0729***** (0.0163)	**0.0731***** (0.0163)
mixed_eth	0.00408 (0.0664)	-0.00719 (0.0662)	-0.00598 (0.0665)
asian_ab	**-0.189*****	**-0.190*****	**-0.191*****

	(0.0388)	(0.0390)	(0.0390)
black_bb	**-0.0807*** (0.0448)	**-0.0857*** (0.0445)	**-0.0859*** (0.0447)
chinese_other	**-0.200**** (0.0882)	**-0.208**** (0.0887)	**-0.204**** (0.0883)
prgamble	**0.0911***** (0.0166)	**0.0922***** (0.0166)	**0.0919***** (0.0166)
ppgamble	**0.115***** (0.0362)	**0.112***** (0.0362)	**0.112***** (0.0363)
drinker	0.00895 (0.0151)	0.0143 (0.0151)	0.0139 (0.0151)
married	0.0255 (0.0170)	0.0278 (0.0174)	**0.0380**** (0.0177)
Wales	-0.0159 (0.0318)	-0.200 (0.0318)	-0.0216 (0.0319)
Scotland	0.0147 (0.0221)	0.0129 (0.0221)	0.0122 (0.0221)
retired	**-0.0608*** (0.0330)	**-0.0825**** (0.0326)	-0.0516 (0.0334)
hdo	**-0.0709***** (0.0330)	**-0.105***** (0.0257)	**-0.0644***** (0.0275)
student	**-0.182***** (0.0402)	**-0.233***** (0.0379)	**-0.174***** (0.0403)
unemployed	0.0358 (0.0449)	-0.0200 (0.0429)	0.0360 (0.0198)
ads	**-0.0356*** (0.0197)	**0.0377*** (0.0197)	**0.0360*** (0.0198)
hsize	-0.0101 (0.00638)	**-0.0245***** (0.00693)	**-0.0275***** (0.00708)
hemployed	0.0229 (0.0246)	0.0192 (0.0251)	0.0340 (0.0254)

* p < 0.10, ** p < 0.05, *** p < 0.01
Robust standard errors in parentheses

Bibliography

1. Amore, MD, Murtinu, S. Tobit models in strategy research: Critical issues and applications. Global Strategy Journal. 2019; 1– 25. [Online] (Accessed 17th January 2020) Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/gsj.1363
2. Anyaegbu, G. (2010). Using the OECD equivalence scale in taxes and benefits analysis. Econ Lab Market Rev 4, 49–54 [Online] (Accessed 5th May 2020) Available at: https://link.springer.com/article/10.1057/elmr.2010.9
3. Beckert, J., & Lutter, M. (2009). The Inequality of Fair Play: Lottery Gambling and Social Stratification in Germany. European Sociological Review, 25(4), 475-488. [Online] (Accessed 15th February 2020) Available at: www.jstor.org/stable/27745226
4. Brunk, G. (1981). A Test of the Friedman-Savage Gambling Model. The Quarterly Journal of Economics, 96(2), 341-348. [Online] (Accessed 24th March 2020) Available at: www.jstor.org/stable/1882394
5. Buddelmeyer, H., Peyton, K. (2014), How Windfall Income Increases Gambling at Poker Machines. Econ Rec, 90: 236-248 [Online] (Accessed 7th April 2020) Available at: https://onlinelibrary.wiley.com/doi/full/10.1111/1475-4932.12094
6. Cameron, A.C., Trivedi, P. K. (2005) “Microeconometrics: Methods and Applications”, Cambridge University Press
7. Castrén, S., Kontto J., Alho, H., Salonen, A.H. (2018) The relationship between gambling expenditure, socio-demographics, health-related correlates and gambling behaviour—a cross-sectional population-based survey in Finland. Addiction, 113: 91– 106. [Online] (Accessed 10th April 2019) Available at: https://onlinelibrary.wiley.com/doi/full/10.1111/add.13929
8. Christiansen, E. (1998). Gambling and the American Economy. The Annals of the American Academy of Political and Social Science, 556, 36-52. [Online] (Accessed 23rd March 2019) Available at: www.jstor.org/stable/1049328
9. Cook, S., Turner, N.E., Ballon, B. et al. (2015). Problem Gambling Among Ontario Students: Associations with Substance Abuse, Mental Health Problems, Suicide Attempts, and Delinquent Behaviours. J Gambl Stud 31, 1121–1134. Online] (Accessed 25th March 2020) Available at: https://link.springer.com/article/10.1007/s10899-014-9483-0#citeas
10. Fong, D., Ozorio, B. (2005). Gambling Participation and Prevalence Estimates of Pathological Gambling in a Far-East Gambling City: Macao. UNLV Gaming Research & Review Journal. 9. 15-28. [Online] (Accessed 10th November 2019) Available at: https://digitalscholarship.unlv.edu/cgi/viewcontent.cgi?article=1158&context=grrj
11. Friedman, M., Savage, L. (1948), The utility analysis of choices involving risk, Journal of Political Economy 56 (4), pp. 279–304. [Online] (Accessed 7th February 2020) Available at: https://www.jstor.org/stable/1826045?seq=1

12. Gambling Commission (2019), Industry Statistics April 2016 to March 2019, United Kingdom. [Online] (Accessed 5th March 2020) Available at: https://www.gamblingcommission.gov.uk/PDF/survey-data/Gambling-industry-statistics.pdf
13. Greene, W. H. (1996). Marginal Effects in the Bivariate Probit Model. NYU Working Paper No. EC-96-11. [Online] (Accessed 12th April 2019) Available at: https://ssrn.com/abstract=1293106
14. Haisley, E., Mostafa R., Loewenstein, G. (2008), Subjective relative income and lottery ticket purchases. J. Behav. Decis. Making, 21: 283-295. [Online] (Accessed 2nd January 2019) Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/bdm.588
15. Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47(2), 263-291. [Online] (Accessed 7th February 2020) Available at: www.jstor.org/stable/1914185
16. McClure, S., Laibson, D.I., Loewenstein, G., Cohen, J.D. (2004). Separate Neural Systems Value Immediate and Delayed Monetary Rewards. Science, 306(5695), 503-507. [Online] (Accessed 7th April 2020) Available at: www.jstor.org/stable/3839330
17. National Centre for Social Research (2010), British Gambling Prevalence Survey, 2010, UK Data Service. [Online] (Accessed 10th November 2019) Available at: https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=6843
18. 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, Vol. 13, Issue 3, pp.257-271 [Online] (Accessed 15th February 2020) Available at: https://www.tandfonline.com/doi/full/10.1080/13668803.2010.488101
19. Shafir, E. LeBoeuf, R.A. (2002). Rationality. Annual Review of Psychology. 53 (1): 491–517. [Online] (Accessed 2nd January 2019) Available at: https://www.annualreviews.org/doi/10.1146/annurev.psych.53.100901.135213
20. Thaler, R. (2015). Misbehaving: The Making of Behavioral Economics. W. W. Norton & Company