

School of Economics Working Paper
2021-01



**SCHOOL OF
ECONOMICS**

Gender and Gambling Preference

Philomena Bacon*

Anna Conte**

Peter G Moffatt*

*University of East Anglia. Norwich

**Department of Statistical Sciences,
Sapienza University of Rome

School of Economics
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
United Kingdom
www.uea.ac.uk/economics

Gender and gambling preference

Philomena Bacon^{a,*}, Anna Conte^{b,†}, Peter G. Moffatt^{a,‡}

^a*School of Economics, University of East Anglia, Norwich, UK*

^b*Department of Statistical Sciences, Sapienza University of Rome, IT*

Abstract

Gambling preferences are analysed using survey data from the wider population. Respondents were confronted with a hypothetical lottery question, in which they were asked to imagine having just won a large prize, and asked how much of this prize they would be willing to invest in a further gamble. We observe the majority of respondents avoiding the gamble altogether. We demonstrate that such behaviour cannot easily be explained by standard models of choice under risk, since it implies implausible degrees of risk aversion. We propose that the observed behaviour can instead be explained in terms of gambling aversion. Since the decision variable takes the form of the number of “units” of the prize that the respondent wishes to invest in the gamble, and since the decision is observed twice for some respondents, we adopt the panel version of the Zero-Inflated Poisson model as an econometric framework. We assume that individual characteristics affect both stages of the decision-making process. We are particularly interested in the effect of gender, and we find that males have a significantly higher probability of participating in the gamble, and are also (conditional on gambling) prepared to gamble significantly larger amounts.

Keywords: Gambling behaviour; Gender; Panel Zero-Inflated Poisson model; survey data

JEL Classification: C33; D01; J16

Acknowledgments: We acknowledge access to the German Socio-Economic Panel (SOEP) for use in this research under licence number: 2596.

Declarations: Codes are available from the authors upon request. The authors declare that they have nothing to disclose.

*Registry and Council House, Norwich Research Park, Norwich NR4 7TJ, UK *Email:* p.bacon@uea.ac.uk

†Viale Regina Elena 295G, 00161 Rome, Italy, tel.: +39 06 49255170, *Email:* anna.conte@uniroma1.it

‡Registry and Council House, Norwich Research Park, Norwich NR4 7TJ, UK *Email:* p.moffatt@uea.ac.uk

1 Introduction

An understanding of gambling preference is important in tackling the social issue of ‘problem gambling’ (Meyer et al., 2009), and also because it plays an important role in certain models in Financial Economics (see e.g. Doran et al., 2012). Empirical research on the determinants of gambling preference often focuses on problem gamblers (LaPlante et al., 2006; Odlaug et al., 2011; Sancho et al., 2019). However, problem gamblers are known to comprise a very small proportion of the population,¹ and hence using such data sources clearly requires adjustments for sample selection. This paper avoids such selection issues by using survey data from the wider population. Another problem associated with surveys targeted at problem gamblers is that such individuals have a tendency to lie when responding to questions relating directly to problem gambling (Johnson et al., 1998) and this brings into questions the validity of the data. In this paper we avoid this problem by analysing survey questions that are not directly posed for the purpose of uncovering problem gambling.

We are particularly interested in the effect of gender on gambling preference. The key questions we aim to answer are which gender is more likely to gamble, and which gender gambles larger amounts. Note that the answers to these two questions are not necessarily the same. The survey response on which we focus is a hypothetical lottery question with high stakes. Survey questions of this type have been useful in estimating the risk attitudes of respondents (see e.g. Bacon et al., 2020). However, such survey questions tend to produce choices that imply implausibly high levels of risk aversion. In this paper, we postulate that, when an individual is faced with a high-stakes lottery, ‘gambling preference’ takes precedence over risk preference in determining such choices. In particular, the high incidence of the choice of the safest alternative may be attributed to a form of ‘gambling aversion’.

We use the same hypothetical lottery choice data used by Bacon et al. (2020), from the German Socio-Economic Panel Survey (SOEP), but we shift the focus from risk preference to gambling preference. The hypothetical question analysed places respondents in the imaginary scenario of having just won a large sum of money, and the question essentially asks how much of the prize that they have just won, they are willing to risk in a further gamble.

¹Epidemiological studies estimate the prevalence of lifetime pathological gambling as 0.4% to 1.5% among adults in the USA (Odlaug et al., 2011). From a total of 202 studies conducted in different countries between 1975 and 2012, Williams et al. (2012) calculate that year rate of problem gambling ranges from 0.5% to 7.6%, with an average rate across all countries of 2.3%.

Hence the setting bears similarities to some settings that have been studied in the gambling addiction literature. For example, in the context of scratch card gambling, Griffiths (1995) finds evidence that receiving winnings acts as a reinforcement to winners to continue gambling, and further notes that the industry exploits this by ensuring that winnings can be gambled immediately. We are particularly interested in this concept of reinforcement. In our context, as already noted, a significant proportion of the sample avoid the gamble, but some do gamble to varying degrees, and this may be partly explained in terms of reinforcement brought about by the large (hypothetical) win.

Since our survey question essentially asks the respondent how many units of €20,000 they would gamble having just won €100,000, we treat the response as a count variable, and apply models appropriate for count data, namely the Poisson regression model. Count data models have been used previously in gambling research by Kastirke et al. (2015). Moreover, in recognition that a proportion of the population are ‘gambling averse’, we adopt a zero-inflated version of the Poisson model (ZIP) (Lambert, 1992), in order to separate out the effects of individual characteristics on participation in gambling and the extent of the gamble. Actually, since some respondents are observed on more than one occasion, we use the panel-data extension of the ZIP model (Wang et al., 2002). Similar models have been used previously in gambling research. For example, Jaunky and Ramchurn (2014) apply a double-hurdle model to scratch card purchases, and find interestingly that while males are less likely to participate than females, participating males tend to spend more than participating females.

There is some evidence on the effect of gender on risk-taking more generally. Dohmen et al. (2011, p. 535) find that males are significantly more willing to take risks than females and that this difference is most pronounced in the domain of financial risk taking. Halko et al. (2012) analyse data from a context-specific cohort of Finnish investors who were given hypothetical investment questions similar to the one we analyse. They find that females are more risk averse but the difference is not statistically significant. Welte et al. (2007) provide interesting evidence on the types of gambling that present the most severe problems for each gender. For males, the problem types are casinos and card games; for females, they are lotteries, casinos, gambling machines and bingo.

The objective of this paper is to estimate the effects of respondent characteristics including gender

on both participation and extent of the gamble, using the aforementioned hypothetical response data from a large survey of the wider population.

The paper is structured as follows. Section 2 describes the data set, Section 3 describes the model, Section 4 presents the results, and Section 5 concludes.

2 Data description

The data set used in this study is extracted from the German Socio-Economic Panel Survey (SOEP) (Wagner et al., 1993, 2007). This is a representative panel survey of the resident adult population of Germany surveying approximately 20,000 households annually, and inquiring into lifestyle and economic activities. It has been running since 1984. In 2004, the survey broadened to include questions associated with risk attitude.

The question on which we focus is the “hypothetical lottery question”. This question was asked in two years of the survey, 2004 and 2009. It takes the following form:

Imagine that you have won €100,000 in the lottery. Immediately after receiving your winnings you receive the following offer: You have the chance to double your money. But it is equally possible that you will lose half the amount invested. You can participate by staking all or part of your €100,000 on the lottery, or choose not to participate at all. What portion of your lottery winnings are you prepared to stake on this financially risky, yet potentially lucrative lottery investment?

€100,000 (i.e. all of it);

€80,000;

€60,000;

€40,000;

€20,000;

Nothing: I would decline the offer

Data on responses to this hypothetical lottery question form the focus of our model of gambling preference which is described in Section 3. From the SOEP data we also extract a number

of demographic variables which are used as explanatory variables in the model. These include gender, age, marital status, and years of education. Descriptive statistics of all variables used, along with SOEP dataset codes, are supplied in Table A.1 of the Appendix.

3 Econometric model

For the purpose of our econometric model of gambling preference, we imagine that, when faced with the hypothetical lottery question described in Section 2, an individual is endowed with five “coupons” worth €20,000 each, and their choice problem amounts to deciding how many of these five “coupons”, if any, to invest in the gamble.² Accordingly we define the dependent variable y_{it} for subject i , $i = 1, \dots, n$, in period $t \in \{2004, 2009\}$, as the number of coupons invested, so that $y_{it} \in \{0, 1, \dots, 5\}$. Hence we treat y_{it} as a count variable, and model it as such.

In our sample, there are 41,704 observations from 27,474 respondents; 14,230 respondents are observed in both years, 7,025 in 2004 only and 6,219 in 2009 only. Of those who are observed in both years, 8,182 (57.5%) choose the same investment both times, 4,252 (29.9%) invested less in 2009 than they did in 2004, 1,796 (12.6%) invested more in 2009. This amounts to strong evidence that investments were lower in 2009 (Wilcoxon signed-rank test: $z = 31.523$, $p\text{-value} = 0.0000$). Bacon et al. (2020) explain this change in terms of risk vulnerability (Gollier and Pratt, 1996). Specifically, agents are more risk averse at times of high background risk, as was the case in 2009, being the nadir of the global recession that followed the global financial crisis.

The distributions of investment by year and gender are shown in Figure 1. In all panels, we see a clear mode at the choice of a zero investment. The proportion of respondents choosing a zero investment is greater than 50% in all cases. However, this mode appears much more prominent in 2009 than in 2004 for both males and females. It is again reasonable to hypothesise that this difference is a consequence of the global financial crisis occurring between the two dates.

In Table A.1 of the Appendix, we report the mean and variance of the number of coupons invested. We see that the variance is somewhat higher than the mean in both years, implying overdispersion. This is a well-known feature of count data, which essentially indicates that the simple Poisson model is inappropriate as a modelling framework. One common cause of overdis-

²For a similar interpretation in a different context, see Conte and Moffatt (2014).



FIGURE 1
Distribution of response to hypothetical lottery question by year and gender

person is the presence of excess zeros (Winkelmann, 2008). It is therefore not surprising that the overdispersion (measured as the variance-mean ratio) appears to be higher in the year 2009, when, as seen clearly in Figure 1, excess zeros were more prominent. The clear presence of excess zeros, and consequent overdispersion, is one of the important data features that guides our choice of econometric modelling strategy.

In Figure 1 we also see evidence that males are less likely than females to choose a zero investment, and that positive investments tend to be higher for males. Note that these two observations provide a non-parametric preview of the key results we will be looking for in the parametric model.

Bacon et al. (2020) have used the same data set to estimate the distribution of risk aversion over the population. They derive ranges of absolute risk aversion corresponding to each choice. For example, they find that, under the assumption of expected utility maximisation, the choice to gamble zero implies a coefficient of absolute risk aversion greater than 4.70. It may be seen as implausible that the majority of the population are this risk averse, not least because it is so at-odds with estimates of risk aversion estimated in other contexts in previous literature.³ For this reason we propose that behaviour in large-stakes hypothetical lotteries may be better modelled in terms of gambling preference rather than risk preference. In particular, we hypothesise that the

³For example, Holt and Laury (2002) report a range of (relative) risk aversion of 0.3-0.5 based on data obtained using multiple price lists, and they cite other literature that has produced similar estimates.

decision to gamble zero is the result of an aversion to any sort of gamble rather than that of very high risk aversion.

Some respondents choosing to invest zero may be labelled as gambling-averse. For those who choose to invest, the number of coupons invested may be interpreted as the level of affinity towards this particular gamble. Of course, an investment of zero does not imply gambling aversion, because the zero investment may be a result of low affinity towards the gamble. Hence there are two types of zero: gambling aversion, and low affinity. Respondents appearing in the right tail of the distributions in Figure 1 have the highest affinity towards the gamble. In order to capture the concepts of both gambling-aversion and affinity towards the gamble, a two-part econometric model is required, with one equation representing the decision to participate in the gamble, and the other representing the level of investment. Note that affinity towards the gamble, represented by the second of these two equations, is closely related to the phenomenon of reinforcement (Griffiths, 1995) that was discussed in Section 1 in the context of scratchcard gambling. Here, we perceive a high investment as a manifestation of reinforcement: the individual is being lured by the experience of the windfall gain into risking a high proportion of the gain in a further gamble.

3.1 The random-effects zero-inflated Poisson model

We model this situation via a random-effects zero-inflated Poisson model (see Crepon and Duguet, 1997, Wang et al., 2002, and Song et al., 2018, among others).⁴

Let p_{it} be the probability of being a gambler in period t , and, consequently, $(1 - p_{it})$ captures the probability of those who are not likely to engage in gambling activities in the current circumstances.

Then, the probability of observing $y_{it} = k$ for subject i in period t is:

$$f(y_{it}; p_{it}, \lambda_{it}) = \begin{cases} (1 - p_{it}) + p_{it} \mathcal{P}(0; \lambda_{it}) & \text{if } k = 0 \\ p_{it} \mathcal{P}(y_{it}; \lambda_{it}) & \text{if } k = 1, \dots, 5 \end{cases} \quad (1)$$

Here, $\mathcal{P}(k; \lambda_{it})$ is the Poisson probability mass function evaluated at k , with parameter $\lambda_{it} > 0$.

⁴Had the hypothetical lottery question that we analyse here been asked in many more waves of the survey so that a much larger span of time had been covered and subjects had been observed repeatedly, we could have attempted to isolate via a hurdle model the ‘unbending non-gamblers’, that is those who would not engage in gambling activities under any circumstances.

Disregarding t , this model is a zero-inflated Poisson model (Lambert, 1992). The idea behind it is that each subject's decision can be regarded as a draw from a Poisson distribution with an abnormal mass at zero. With probability $(1 - p_{it})$, the subject is a non-gambler and falls in the excess mass at zero of the distribution; with probability p_{it} , the subject's decision is driven from a standard Poisson distribution, that is they can decide to invest any number of coupons from the amount previously won. This number includes zero, given that they might be gamblers but unwilling to invest in this particular lottery.

We said that we distinguish between gamblers and non-gamblers in the current circumstances. In some sense, we need to control for those who do not gamble in order to be able to identify the features of those who are prone to gambling addiction, that we identify as those who are in the upper tail of the Poisson distribution. The probability of being a gambler, p_{it} , is modelled via a normal probability distribution function, $\Phi(\cdot)$, and depends on the vector of individual characteristics (including an intercept) intended to capture the "current circumstances", x_{it} , in the following way:

$$p_{it} = \Phi(x'_{it}\gamma + \delta_i) \quad (2)$$

Here, γ is a vector of coefficients on the variables in x_{it} to be estimated and δ_i is an individual-specific intercept. Basically, this component of the model is similar to a probit model. If a subject is a non-gambler, based on the current circumstances, they should score a zero and contribute to the observed mass at zero. Otherwise, they are a gambler and decide on an amount (taken from the amount previously won) to invest in the proposed lottery.

The intensity of the attitude to gambling, signalled by the number of €20,000 notes invested, depends on the parameter λ_{it} according to some individual characteristics, subsumed by the vector x_{it} ,⁵ and an individual-specific term α_i as follows:

$$\lambda_{it} = \exp(x'_{it}\beta + \alpha_i) \quad (3)$$

⁵We note that the vector of individual characteristics x_{it} needs not to be the same in the equations for λ_{it} and p_{it} , but there is no practical reason, if not theoretical, that prevents this.

where β is a vector of coefficients to be estimated.

We assume that the two individual-specific random variables α_i and δ_i follow a bivariate normal distribution:

$$\begin{pmatrix} \alpha_i \\ \delta_i \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\delta \\ \rho\sigma_\alpha\sigma_\delta & \sigma_\delta^2 \end{pmatrix} \right) \quad (4)$$

where σ represents the standard deviations and ρ the correlation.

The probability of observing $y_{it} = k$ is

$$\begin{aligned} \Pr(y_{it} = k | \alpha_i, \delta_i) &= \{[1 - \Phi(x'_{it}\gamma + \delta_i)] + \Phi(x'_{it}\gamma + \delta_i) \times \mathcal{P}(0; x'_{it}\beta + \alpha_i)\}^{\mathbf{1}_{[y_{it}=0]}} \\ &\quad + \{\Phi(x'_{it}\gamma + \delta_i) \times \mathcal{P}(k; x'_{it}\beta + \alpha_i)\}^{\mathbf{1}_{[y_{it} \neq 0]}} \end{aligned} \quad (5)$$

Assuming that, conditional on α_i and δ_i , the probability of observing a certain outcome is independent over individual-specific random variables and everything else in the model, we can write the individual likelihood contribution for subject i observed in the two periods $t = 1, 2$ as :

$$\begin{aligned} \Pr(y_{i1}, y_{i2} | \alpha_i, \delta_i) &= \prod_{t=1,2} \left\{ \{[1 - \Phi(x'_{it}\gamma + \delta_i)] + \Phi(x'_{it}\gamma + \delta_i) \times \mathcal{P}(0; \exp(x'_{it}\beta + \alpha_i))\}^{\mathbf{1}_{[y_{it}=0]}} \right. \\ &\quad \left. + \{\Phi(x'_{it}\gamma + \delta_i) \times \mathcal{P}(y_{it}; \exp(x'_{it}\beta + \alpha_i))\}^{\mathbf{1}_{[y_{it} \neq 0]}} \right\} \end{aligned} \quad (6)$$

By integrating out α_i and δ_i , we obtain the unconditional individual likelihood contribution as:

$$\Pr(y_{i1}, y_{i2}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Pr(y_{i1}, y_{i2} | \alpha, \delta) g(\alpha, \delta) d\alpha d\delta \quad (7)$$

where $g(a, b)$ is the bivariate normal density function evaluated at a and b .

The full-sample log-likelihood for the sample of individuals in the two periods is

$$\text{Log-Likelihood} = \sum_{i=1}^n \log \Pr(y_{i1}, y_{i2}) \quad (8)$$

4 Estimation results

The model is estimated using data from the full sample and separately from the male and female subsamples. To estimate the model, we use the method of Maximum Simulated Likelihood. The double integral appearing in (7) is evaluated using two sets of Halton sequences (100 draws per subject).

For the sake of checking the robustness of our main results, we estimate three different specifications of the model. Table 1 presents the results for specification 1. This includes basic characteristics of the individual and also a dummy for the second survey year 2009. Table 2 presents the results for specification 2 which adds a control for risk-lovingness⁶ to specification 1. Table 3 shows the results for specification 3 which adds a set of self-reported life-satisfaction variables (all in binary form) to specification 1. For each (sub)sample, results in the tables are displayed in two columns: one, labelled ‘Participation’, relates to the probability (p_{it} defined in (2)) of participating in the gamble under the prevailing circumstances captured by the explanatory variables; the other, labelled ‘Investment’, shows the factors that influence the amount invested ($\log(\lambda_{it})$ defined in (3)) measured in number of €20,000 “coupons”.

The key results are similar between the three specifications, so we shall commence by focusing on specification 1. Our primary interest is in the effect of gender. We see that being a male has a strongly significant effect on both the probability of participation in the gamble, and the amount invested. From the results of specification 1 (Table 1), we have computed marginal effects.⁷ These tell us that males are 5 percentage points more likely to participate, and that participating males invest 0.46 more coupons than participating females. Overall, males invest 0.28 coupons more than females. Having established the strong gender effect in both equations, we shall consider the

⁶The risk-lovingness variable is obtained from the response to a question of the form “Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?”. The response was measured on a 0-10 Likert Scale.

⁷The marginal effect of male on probability of participation is 0.05177 (std. err. 0.00005); the marginal effect of male on count conditional on participation is 0.45728 (std. err. 0.00260); the unconditional effect of male on count is 0.276523 (std. err. 0.00013). These figures are calculated for specification 1, full sample in Table 1

TABLE 1
Estimation results (specification 1)

	PARTICIPATION (p_{it})			INVESTMENT ($\log(\lambda_{it})$)		
	full sample	male	female	full sample	male	female
male	0.155*** (0.031)	—	—	0.254*** (0.019)	—	—
year 2009	-0.725*** (0.037)	-0.629*** (0.046)	-0.857*** (0.064)	-0.043** (0.020)	-0.040 (0.025)	-0.037 (0.033)
age	-0.022*** (0.002)	-0.019*** (0.002)	-0.026*** (0.003)	-0.001 (0.001)	-0.003*** (0.001)	0.002 (0.002)
years of education	0.066*** (0.005)	0.058*** (0.007)	0.080*** (0.009)	0.006** (0.003)	0.015*** (0.004)	-0.006 (0.005)
married	-0.169*** (0.052)	-0.102 (0.069)	-0.261*** (0.084)	-0.105*** (0.031)	-0.089** (0.041)	-0.114** (0.050)
separated	0.209* (0.125)	0.008 (0.151)	0.550** (0.236)	-0.058 (0.071)	0.009 (0.090)	-0.187 (0.115)
divorced	-0.060 (0.070)	0.045 (0.097)	-0.168 (0.108)	-0.070 (0.044)	-0.071 (0.059)	-0.061 (0.070)
widowed	-0.098 (0.083)	-0.084 (0.137)	-0.099 (0.119)	0.114* (0.061)	0.030 (0.103)	0.068 (0.085)
# children in HH	-0.044** (0.020)	-0.015 (0.026)	-0.086** (0.033)	0.010 (0.012)	0.013 (0.015)	0.007 (0.020)
log net HH income	0.225*** (0.026)	0.224*** (0.035)	0.249*** (0.041)	0.048*** (0.015)	0.249*** (0.041)	0.010 (0.024)
employed	0.064* (0.036)	0.065 (0.049)	0.051 (0.056)	0.016 (0.023)	0.003 (0.030)	0.020 (0.035)
home owner	0.170*** (0.035)	0.161*** (0.046)	0.184*** (0.055)	-0.096*** (0.023)	-0.077*** (0.029)	-0.117*** (0.036)
constant	-1.889*** (0.264)	-1.845*** (0.356)	-1.937*** (0.413)	-0.478*** (0.160)	-0.543** (0.210)	-0.043 (0.252)
σ_δ				0.967*** (0.063)	0.942*** (0.081)	1.057*** (0.103)
σ_α				0.356*** (0.018)	0.328*** (0.023)	0.397*** (0.028)
ρ				0.366*** (0.116)	0.502*** (0.167)	0.114 (0.140)
Log-likelihood				-40715.052	-21655.546	-19015.221
N observations				41,704	19,996	21,699
N subjects				27,474	13,265	14,209

Note: Standard errors in parentheses.

* p -value < 0.10, ** p -value < 0.05, *** p -value < 0.01

estimates from the male and female subsamples separately in further interpretations.

The year effect appears to be important for both genders. In 2009, individuals of both genders are significantly less likely to participate in the gamble. The effect of year 2009 on amount gambled is also negative, but much weaker. These results are consistent with the observations made when examining Figure 1, where we interpreted differences between 2004 and 2009 in terms of the effect of the 2007-2008 Global Financial Crisis on gambling preferences. The effect of the crisis increasing

gambling aversion, while not significantly influencing the amount gambled, is just the sort of result that underlines the importance of the 2-equation approach adopted in this paper.

Our results reveal that gambling is less popular among older people of either gender, but the amount invested decreases with age only for males. These results are in broad agreement with those of Brochadoa et al. (2018). We have estimated alternative specifications with powers of ‘age’ in order to capture possible non-linearities in the effect of age on whether and how much to gamble, but we found no statistically significant non-linearities.⁸ It must be said that evidence of the effect of age on gambling behaviour is very mixed. Interestingly, some researchers have found evidence of a U-shaped effect of age, with an increase in gambling later in life being attributed to increased leisure time and fewer financial commitments. According to Van der Maas et al. (2017), gambling is an attractive leisure activity for older adults for several reasons: gambling can offer opportunities to socialize, to participate in economic activity and to take risks that may otherwise be less accessible to them. Epidemiological research stresses that the individuals’ chronological age and sex are crucial demographic variables for predicting gambling preference (Jiménez-Murcia et al., 2020). Welte et al. (2007) find that lottery gambling is much more of a factor in gambling problems for respondents 30 years or older, whereas Jaunky and Ramchurn (2014) do not find evidence of an age related effect. Rutledge et al. (2016) focus on age and gambling preferences in a study involving more than 25,000 participants across the age spectrum using a smartphone experiment. The authors find a correlation with a decline in dopamine levels which naturally falls with age. Patients who receive supplements of this drug to treat certain conditions, such as Parkinson’s, have been reported to exhibit lower levels of risk aversion and higher instances of gambling disorder. Balabanis (2001), on the other-hand, claims that ‘Gender and age were not found to be related with buying compulsiveness in lottery ticket and scratch-cards’. He finds that buying lottery tickets and scratch-cards are closely interrelated and associated with compulsiveness, and impulse behaviour. He postulates that because of the common distribution channels, there is a strong correlation between the purchase of lottery tickets, scratch-cards and heavy smokers.⁹

In our models, marital status is controlled for via five dummies: single (base case), married, separated, divorced and widowed. We find that for females, being married has a negative effect

⁸The results are available from the authors upon request.

⁹We could not include an indicator for smoking habits, since this information was available for only one of the survey years we consider.

TABLE 2
Estimation results (specification 2)

	PARTICIPATION (p_{it})			INVESTMENT ($\log(\lambda_{it})$)		
	full sample	male	female	full sample	male	female
male	0.121*** (0.031)			0.167*** (0.019)		
year 2009	-0.714*** (0.037)	-0.601*** (0.046)	-0.852*** (0.064)	0.002 (0.019)	0.008 (0.024)	-0.006 (0.032)
age	-0.021*** (0.002)	-0.018*** (0.002)	-0.025*** (0.003)	0.000 (0.001)	-0.001 (0.001)	0.003* (0.001)
years of education	0.063*** (0.005)	0.051*** (0.007)	0.079*** (0.009)	0.003 (0.003)	0.010** (0.004)	-0.006 (0.005)
married	-0.152*** (0.052)	-0.081 (0.068)	-0.249*** (0.083)	-0.088*** (0.031)	-0.079** (0.039)	-0.091* (0.049)
separated	0.221* (0.125)	0.017 (0.151)	0.535** (0.229)	-0.072 (0.069)	-0.004 (0.088)	-0.179 (0.112)
divorced	-0.048 (0.069)	0.052 (0.096)	-0.150 (0.106)	-0.078* (0.043)	-0.093 (0.057)	-0.058 (0.068)
widowed	-0.092 (0.082)	-0.086 (0.134)	-0.080 (0.118)	0.105* (0.060)	0.069 (0.101)	0.056 (0.083)
# children in HH	-0.043** (0.020)	-0.014 (0.026)	-0.085** (0.033)	0.015 (0.011)	0.017 (0.014)	0.013 (0.019)
log net HH income	0.212*** (0.026)	0.194*** (0.035)	0.249*** (0.040)	0.017 (0.015)	0.039** (0.019)	-0.010 (0.023)
employed	0.061* (0.036)	0.067 (0.048)	0.045 (0.056)	0.010 (0.022)	-0.003 (0.029)	0.018 (0.034)
home owner	0.171*** (0.035)	0.159*** (0.046)	0.186*** (0.055)	-0.112*** (0.022)	-0.099*** (0.028)	-0.125*** (0.035)
risk loving	0.562*** (0.039)	0.602*** (0.050)	0.510*** (0.065)	0.439*** (0.020)	0.433*** (0.024)	0.434*** (0.033)
constant	-1.867*** (0.262)	-1.699*** (0.351)	-2.038*** (0.408)	-0.228 (0.156)	-0.263 (0.204)	0.048 (0.245)
σ_δ				0.945*** (0.064)	0.907*** (0.081)	1.030*** (0.103)
σ_α				0.287*** (0.020)	0.241*** (0.027)	0.348*** (0.030)
ρ				0.306** (0.141)	0.420* (0.225)	0.129 (0.164)
Log-likelihood				-40038,954	-21210.45	-18790.322
number of observations				41,695	19,996	21,699
number of subjects				27,474	13,265	14,209

Note: Standard errors in parentheses.

* p -value < 0.10, ** p -value < 0.05, *** p -value < 0.01

on both participation and investment. For males, there is only an effect on investment and it is smaller than that for females (cf. Carneiro et al., 2020). Being separated appears to make females (but not males) more likely to participate. This result may be linked to research by Brochadoa et al. (2018) who claim that lotteries have become a leisure activity and a refuge for women from a ‘sense of alienation’. The presence of children has a negative effect on female participation but no effect on males. These interesting gender differences can be linked to the findings of Casey (2006) who argues that ‘as a leisure and gambling activity, which has attracted women players in

unprecedented numbers, National Lottery play constitutes an important part of many women’s everyday lives and that women’s class and gender identities were formed, reproduced and developed through their National Lottery play’.

In our estimations, both years of education and household income tend to have a positive effect on participation, and for males only, investment. These findings appear to contradict some previous findings. Jiménez-Murcia et al. (2020) reported that characteristics such as lower education levels and lower socio-economic status predict higher interest in non-strategic gambling. Haisley et al. (2008) find that “lotteries are more alluring for poor people because they provide an opportunity to correct for low-income status”. This difference in findings is surely explained by the types of lottery being considered. It is reasonable to expect low-income and poorly educated individuals to be attracted to low-stakes lotteries of the type studied in the previous literature just cited, but to avoid lotteries with considerably higher stakes of the type considered here.

Interestingly, we see that for both males and females, home-ownership tends to increase the participation probability, but to reduce the amount invested. Perhaps those who own a home feel a sense of economic stability that allows them to indulge in gambling activities, but expenses relating to home-ownership such as mortgage payments and utility bills restrain their appetite for gambling.

The results presented in Table 2, which adds to specification 1 a dummy variable taking the value 1 if the individual reports to be risk loving, 0 otherwise, and Table 3, which includes some attitudinal variables, are very similar to those of specification 1 just discussed. Regarding the additional variables, we see in Table 2 that “risk-lovingness” variable has a strongly positive effect in both equations, for both genders. This result amounts to a useful check of data validation. Table 3 tells us that some of the attitudinal variables have significant effects. One result that is particularly striking is that being “worried about job” increases the probability of participation (for males at least), but reduces the amount invested. A potential problem is that risk-lovingness and/or some of the attitudinal variables included may be tainted by endogeneity. Notwithstanding this problem we consider the reporting of the results from specifications 2 and 3 as useful as robustness checks.

One final gender difference seen in all specifications is that the estimate of the parameter ρ (the correlation between the random effect terms of the two equations) is strongly positive for males, but insignificant for females. The interpretation of this result is that a male who is unusually likely to participate in the gamble may also be expected to invest a large amount, while the same

may not be said for females. An explanation of this difference is left for further research.

5 Conclusion

Agents' reluctance to take part in highly favourable gambles involving high stakes cannot easily be explained by standard models of choice under risk, since such behaviour can only be explained by implausibly high levels of risk aversion. We have therefore adopted the concept of gambling preference and have set out to explain such behaviour in terms of gambling aversion, which may be perceived as overruling risk attitude in certain settings. Of course, some agents do opt for gambles and we have linked this behaviour to the concept of "reinforcement" discussed in the gambling literature, whereby a large win has the effect of encouraging some agents to continue gambling. The econometric framework we have adopted allows for both gambling aversion and reinforcement, and it has been interesting to investigate how the decisions of whether and how much to gamble depend on individual characteristics, most importantly gender. This econometric exercise has been the focus of this paper.

Based on the structure of the data set and the statistical properties of the dependent variable, we have chosen to use a panel version of the zero-inflated Poisson model. Our key findings are that males have a significantly higher probability of participating in the gamble, and are also (conditional on gambling) prepared to gamble significantly larger amounts. Also, family circumstances appear to influence females' gambling behaviour more than that of males.

Our findings are in broad agreement with results from many studies appearing in the gambling literature, despite major differences in the type of gambling activity being considered and in the sampling frame being used. This can be seen as a validation of the type of experimental task considered here, and justifies the use of such tasks in eliciting hard-to-measure behaviour in surveys of the wider population.

TABLE 3
Estimation results (specification 3)

	PARTICIPATION (p_{it})			INVESTMENT ($\log(\lambda_{it})$)		
	full sample	male	female	full sample	male	female
male	0.160*** (0.034)	—	—	0.248*** (0.020)	—	—
year 2009	-0.732*** (0.043)	-0.618*** (0.048)	-0.852*** (0.070)	-0.037* (0.022)	-0.035 (0.056)	-0.043 (0.035)
age	-0.020*** (0.002)	-0.018*** (0.002)	-0.024*** (0.003)	-0.001 (0.001)	-0.003*** (0.001)	0.002 (0.002)
years of education	0.063*** (0.006)	0.056*** (0.007)	0.076*** (0.010)	0.004 (0.003)	0.013*** (0.004)	-0.007 (0.005)
married	-0.156*** (0.055)	-0.087 (0.071)	-0.263*** (0.087)	-0.095*** (0.032)	-0.074* (0.041)	-0.098* (0.052)
separated	0.287** (0.130)	0.122 (0.157)	0.540** (0.230)	-0.057 (0.072)	0.001 (0.092)	-0.154 (0.116)
divorced	-0.058 (0.073)	0.071 (0.098)	-0.188* (0.111)	-0.052 (0.046)	-0.048 (0.060)	-0.047 (.072)
widowed	-0.093 (0.087)	-0.050 (0.141)	-0.124 (0.122)	0.113* (0.064)	0.037 (0.105)	0.071 (0.088)
# children in HH	-0.036 (0.023)	-0.002 (0.029)	-0.071* (0.037)	0.015 (0.013)	0.022 (0.016)	-0.002 (0.022)
log net HH income	0.170*** (0.029)	0.162*** (0.039)	0.194*** (0.044)	0.056*** (0.017)	0.092*** (0.022)	0.016 (0.026)
employed	0.055 (0.046)	0.055 (0.060)	0.031 (0.070)	0.038 (0.028)	0.016 (0.037)	0.057 (0.043)
home owner	0.149*** (0.037)	0.154*** (0.047)	0.141** (0.057)	-0.087*** (0.023)	-0.071** (0.029)	-0.104*** (0.038)
satisfied with health	0.078** (0.038)	0.061 (0.050)	0.108* (0.058)	0.003 (0.025)	-0.026 (0.032)	0.029 (0.040)
satisfied with work	-0.063 (0.042)	-0.097* (0.054)	0.012 (0.067)	0.017 (0.025)	0.027 (0.032)	-0.022 (0.040)
satisfied with housework	-0.072** (0.033)	-0.090** (0.041)	-0.031 (0.054)	-0.029 (0.020)	-0.013 (0.025)	-0.061* (0.034)
satisfied with HH income	0.122*** (0.039)	0.105** (0.051)	0.133** (0.061)	-0.030 (0.025)	-0.023 (0.032)	-0.036 (0.040)
satisfied with dwelling	0.041 (0.049)	-0.017 (0.065)	0.121 (0.076)	-0.118*** (0.031)	-0.127*** (0.040)	-0.129** (0.050)
satisfied with leisure	0.051 (0.038)	0.042 (0.049)	0.035 (0.062)	-0.096*** (0.023)	-0.092*** (0.029)	-0.082** (0.039)
satisfied with childcare	-0.062 (0.056)	-0.058 (0.072)	-0.090 (0.089)	-0.059* (0.033)	-0.069* (0.041)	-0.044 (0.054)
satisfied with life in 5 years	0.239*** (0.043)	0.272*** (0.055)	0.193*** (0.065)	0.101*** (0.029)	0.129*** (0.037)	0.070 (0.047)
worried about finance	-0.042 (0.039)	-0.049 (0.050)	-0.031 (0.061)	-0.046* (0.023)	-0.036 (0.030)	-0.059 (0.039)
worried about job	0.097** (0.040)	0.137*** (0.050)	0.003 (0.065)	-0.103*** (0.023)	-0.110* (0.029)	-0.073*** (0.039)
constant	-1.671*** (0.305)	-1.518*** (0.400)	-1.780*** (0.460)	-0.328* (0.187)	-0.465* (0.239)	0.136 (0.287)
σ_δ				0.988*** (0.081)	0.926*** (0.089)	1.017*** (0.124)
σ_α				0.334*** (0.019)	0.289*** (0.026)	0.383*** (0.030)
ρ				0.299* (0.168)	0.535** (0.227)	0.180 (0.206)
Log-likelihood				-37488.481	-19985.163	-17456.091
N observations				38,355	18,431	19,924
N subjects				25,989	12,562	13,427

Note: Standard errors in parentheses.

* p -value < 0.10, ** p -value < 0.05, *** p -value < 0.01

References

- Bacon, P., Conte, A., and Moffatt, P. G. (2020). A test of risk vulnerability in the wider population. *Theory Decision*, 88:37–50.
- Balabanis, G. (2001). The relationship between lottery ticket and scratch-card buying behaviour, personality and other compulsive behaviours. *Journal of Consumer Behaviour*, 2(1):7–22.
- Brochadao, A., Santosb, M., Oliveirac, F., and Esperançab, J. (2018). Gambling behavior: Instant versus traditional lotteries. *Journal of Business Research*.
- Carneiro, E., Tavares, H., Sanches, M., Pinsky, I., Caetano, R., Zaleski, M., and Laranjeira, R. (2020). Gender differences in gambling exposure and at-risk gambling behavior. *Journal of Gambling Studies*, 36:445–457.
- Casey, E. (2006). Domesticating gambling: Gender, caring and the UK National Lottery. *Leisure Studies*, 25:3–16.
- Conte, A. and Moffatt, P. (2014). The econometric modelling of social preferences. *Theory and Decision*, 76(1):119–145.
- Crepon, B. and Duguet, E. (1997). Research and development, competition and innovation pseudo-maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity. *Journal of Econometrics*, 79(2):355–378.
- Dohmen, T. J., Falk, A., Huffman, D., Schupp, J., and Wagner, G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.
- Doran, J. S., Jiang, D., and Peterson, D. R. (2012). Gambling preference and the New Year effect of assets with lottery features. *Review of Finance*, 16(3):685–731.
- Gollier, C. and Pratt, J. W. (1996). Risk vulnerability and the tempering effect of background risk. *Econometrica*, 64(5):1109–1123.
- Griffiths, M. D. (1995). Scratch-card gambling: A potential addiction? *Education and Health*, 13(2):17–20.
- Haisley, E., Mostafa, R., and Loewenstein, G. (2008). Subjective relative income and lottery ticket purchases. *Journal of Behavioral Decision Making*, 21:283–295.
- Halko, M.-L., Kaustia, M., and Alanko, E. (2012). The gender effect in risky asset holdings. *Journal of Economic Behavior & Organization*, 83:66–81.
- Holt, C. A. and Laury, S. (2002). Risk aversion and incentive effects. *American Economic Review*, 92:1644–1655.
- Jaunky, C. V. and Ramchurn, B. (2014). Consumer behaviour in the scratch card market: a double-hurdle approach. *International Gambling Studies*, 14(1):96–114.
- Jiménez-Murcia, S., Granero, R., Fernández-Aranda, F., and Menchón, J. M. (2020). Comparison of gambling profiles based on strategic versus non-strategic preferences. *Current Opinion in Behavioral Sciences*, 31:13 – 20.
- Johnson, E. E., Hamer, R. M., and Nora, R. M. (1998). The lie/bet questionnaire for screening pathological gamblers: A follow-up study. *Psychological Reports*, 83(3-suppl):1219–1224. PMID: 10079719.

- Kastirke, N., Rumpf, H. J., John, U., Bischof, A., and Meyer, C. (2015). Demographic risk factors and gambling preference may not explain the high prevalence of gambling problems among the population with migration background: results from a German nationwide survey. *Journal of gambling studies*, 31(3):741–757.
- Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1):1–14.
- LaPlante, D., Nelson, S., LaBrie, R., and Shaffer, H. J. (2006). Men and women playing games: Gender and the gambling preferences of iowa gambling treatment program participants. *Journal of Gambling Studies*, 22:65.
- Meyer, G., Hayer, T., and Griffiths, M. (2009). *Problem gambling in Europe: Challenges, prevention, and interventions*. Springer Science and Business Media.
- Odlaug, B. L., Marsh, P. J., Kim, S. W., and Grant, J. E. (2011). Strategic vs nonstrategic gambling: characteristics of pathological gamblers based on gambling preference. *Annals of clinical psychiatry: official journal of the American Academy of Clinical Psychiatrists*, 23(2):105.
- Rutledge, R. B., Smittenaar, P., Zeidman, P., Brown, H. R., Adams, R. A., Lindenberg, U., Dayan, P., and Dolan, R. J. (2016). Risk taking for potential reward decreases across the lifespan. *Current Biology Report* 26, 1634–1639, 26:1634–1639.
- Sancho, M., De Gracia De Gregorio, M., Granero, R., González-Simarro, S., Sánchez, I., Fernández-Aranda, F., and Mena-Moreno, T. (2019). Differences in emotion regulation considering sex, age and gambling preferences in a sample of gambling disorder patients. *Frontiers in psychiatry*, 10:625.
- Song, X. X., Zhao, Q., Tao, T., Zhou, C. M., Diwan, V. K., and Xu, B. (2018). Applying the zero-inflated poisson model with random effects to detect abnormal rises in school absenteeism indicating infectious diseases outbreak. *Epidemiology and Infection*, 146(12):1565–1571.
- Van der Maas, M., Mann, R. E., Matheson, F. I., Turner, N. E., Hamilton, H. A., and McCready, J. (2017). A free ride? an analysis of the association of casino bus tours and problem gambling among older adults. *Addiction*, 112(12):2217–2224.
- Wagner, G. G., Burkhauser, R. V., and Behringer, F. (1993). The English language public use file of the German Socio-Economic Panel. *The Journal of Human Resources*, 28:429–433.
- Wagner, G. G., Frick, J. R., and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP): Scope, Evolution and Enhancements. SOEPpapers on Multidisciplinary Panel Data Research 1, DIW Berlin, The German Socio-Economic Panel (SOEP).
- Wang, K., Yau, K. K. W., and Lee, A. H. (2002). A zero-inflated poisson mixed model to analyze diagnosis related groups with majority of same-day hospital stays. *Computer methods and programs in biomedicine*, 68(3):195–203.
- Welte, J. W., Barnes, G. M., Wiczorek, W. F., Tidwell, M.-C. O., and Hoffman, J. H. (2007). Type of gambling and availability as risk factors for problem gambling: A tobit regression analysis by age and gender. *International Gambling Studies*, 7(2):183–198.
- Williams, R., Volberg, R., and Stevens, R. (2012). The population prevalence of problem gambling: Methodological influences, standardized rates, jurisdictional differences, and worldwide trends. In *Report prepared for the Ontario problem gambling research centre and the Ontario Ministry of Health and Long Term Care*, Ontario: Ontario Problem Gambling Research Centre and the Ontario Ministry of Health and Long Term Care.
- Winkelmann, R. (2008). *Econometric analysis of count data*. Springer Science and Business Media.

Appendix

A Summary statistics

TABLE A.1
Descriptive statistics

Data Code	Description	Statistics	Survey Years	
			2004	2009
plh0203	Risk-win-lottery question	mean	0.71	0.46
		variance	1.17	0.98
d1110211	% Male		48.17	47.74
d11101	Age	mean	47.67	50.02
		sd	16.99	17.56
d11106	Years of education	mean	11.79	11.92
d11104	% Married		61.90	59.73
	% Single		23.23	23.43
	% Separated		13.10	13.35
	% Divorced		7.05	8.07
	% Widowed		6.07	6.95
d11107	# children in HH	mean	0.53	0.44
		max	9	8
up2'j'03 & zp2'j'03	log net HH income	mean	10.36	10.37
		sd	0.64	0.66
up1a01 & up1b01 & zp1a01 & zp1b01	% employed		60.88	59.43
sp85a01 & xp126a01	% home owner		37.80	37.07
up12002 & zp11802	% risk loving		21.27	15.80
plh0171	% satisfied with health		69.56	69.45
plh0173	% satisfied with work		46.94	46.58
plh0174	% satisfied with housework		52.94	58.08
plh0175	% satisfied with HH income		63.94	67.44
plh0177	% satisfied with dwelling		88.05	89.15
plh0178	% satisfied with (amount of) leisure (time)		75.13	77.77
plh0179	% satisfied with child care		12.29	11.27
plh0183	% satisfied with life 5 years (from now)		73.87	77.97
plh0033	% worried about finance		76.10	74.03
plh0042	% worried about job (security)		36.05	31.93
	number of observations (n)		21,251	20,444

Note: First column shows SOEP data codes for each variable.

sd is standard deviation; HH is household.

The variables “risk loving” (obtained from the variable “Willingness To Take Risks In Financial Matters”) and “satisfied with...” are originally measured on a Likert scale 0-10 (low-high), here transformed into binary variables taking the value 0 for categories 0-5 and 1 for 6-10.

The variables “worried about...” are originally reported on a Likert scale 1-3 (very-somewhat-not at all), here transformed into binary variables taking the value 0 for category 3 and 1 for 1-2.

log net HH income is obtained as $\sum_{j \in \{a,b,c,d,f,h,j,k,l,m,o\}} up2'j'03$ in 2004 and $\sum_{j \in \{a,b,c,d,f,h,j,k,l,m,o\}} zp2'j'03$ in 2009.

Information on home ownership is obtained from the 2002 and 2007 waves for year 2004 and 2009, respectively.