UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

CES
Centre d'Économie de la Sorbonne
UMR 8174

# The Importance of Capital in Closing the Entrepreneurial Gender Gap: A Longitudinal Study of Lottery Wins

Sarah FLÈCHE, Anthony LEPINTEUR, Nattavudh POWDTHAVEE

**2021.15**

cnrs
dépasser les frontières

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
https://centredeconomiesorbonne.cnrs.fr/
ISSN : 1955-611X

# The Importance of Capital in Closing the Entrepreneurial Gender Gap: A Longitudinal Study of Lottery Wins

Sarah Flèche[1], Anthony Lepinteur[2] and Nattavudh Powdthavee[3]

## Abstract

Would improving women's access to capital reduce the gender entrepreneurial gap? We study this issue by exploiting longitudinal data on lottery winners. Comparing between large to small winners, we find that an increase in lottery win in period *t*-1 significantly increases the likelihood of becoming self-employed in period *t*. This windfall effect is statistically the same in magnitude for men and women; the top 25% winners (an average win = £831.16) in year *t*-1 report a significant increase in the probability of self-employment in year *t* by approximately 2 percentage points, which is approximately 20-30% of the gender entrepreneurial gap. These results suggest that we can causally reduce the gender entrepreneurial gap by improving women's access to capital that might not be as readily available to the aspiring female entrepreneurs as it is to male entrepreneurs.

Keywords: Gender inequality; self-employment; lottery wins; BHPS
JEL Codes: J16; J21; J24

[1] Centre d'Economie de la Sorbonne - CNRS and CEP, London School of Economics. Email: sarah.fleche@univ-paris1.fr
[2] University of Luxembourg, Maison des Sciences Humaines. Email: anthony.lepinteur@uni.lu
[3] Warwick Business School and IZA. Email: nattavudh.powdthave@wbs.ac.uk
*Acknowledgements:* This work was supported by the French National Research Agency Grant (ANR-17-EURE-0020).

## 1. Introduction

It is well-established that there are many more male than female entrepreneurs in most economically developed countries (Blanchflower and Meyer, 1994; Blanchflower, 2000). For example, according to the Office for National Statistics, around 4.8 million individuals were self-employed in the UK in 2018. Of those, only 33% of them were women. The UK is not an exception: men are substantially more likely to be self-employed than women in all OECD countries.[4] This raises two important questions for our society: What causes this gender entrepreneurial gap to exist in the modern society, and how do we close it?

Currently, there is a large and growing literature on the determinants of entrepreneurship. Previous studies have identified the lack of capital and limited access to the credit market as one of the main barriers to entry for aspiring entrepreneurs (Blanchflower and Oswald, 1998; Evans and Leighton, 1989; Evans and Jovanovic, 1989; Cagetti and De Nardi, 2006; Lindh and Ohlsson, 1996). Other main determinants of self-employment include the earning differentials between full-time employment and self-employment (e.g., Rees and Shah, 1986; Fujii and Hawley, 1991; Taylor, 1996), parental labour force status (Fairlie and Robb, 2007; Djankov *et al.*, 2008; Colombier and Masclet, 2008; Clark and Lepinteur, 2020), the individual's desire for independence, flexibility, as well as the personality traits such as risk-tolerance of the aspiring entrepreneurs (Caliendo *et al.*, 2014; Ardagna and Lusardi, 2008; Cramer *et al.*, 2002). However, relatively less

[4] See https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/trendsinselfemploymentintheuk/2018-02-07 and https://www.oecd-ilibrary.org/employment/entrepreneurship-at-a-glance-2017/gender-differences-in-self-employment-rates_entrepreneur_aag-2017-22-en for more stylised facts about the gender entrepreneurial gap.

research attention has been paid to understanding the causes of gender entrepreneurial gap and, as a result, the process required to reducing it continues to remain imperfectly understood.

One of the most notable contributions in this area comes from a study by Koellinger *et al.* (2013). Using data form 17 developed countries, they find that the lower rate of female business ownership is mainly due to women's lower propensity to start businesses rather than differences in survival rates across genders. Based on a multivariate framework to understand gender differences in self-employment take-up rates, Georgellis and Wall (2005) show that women's decisions to become self-employed are better explained by non-pecuniary aspects of self-employment than economic factors such as liquidity constraints or earning potentials. The authors conclude that wealth is important for men considering self-employment, but not for women. According to these results, a policy aimed at easing the access to the capital market for women would not increase their likelihood of becoming self-employed.[5]

Although women's decisions to enter self-employment are more likely than men's to be determined by non-pecuniary factors – e.g., the cost of childcare (Connelly, 1992), how much self-employment offers them in terms of time flexibility, and the opportunity for them to work from home, the absence of correlation between wealth and self-employment for women comes at odds with other findings. Keeping a wide range of characteristics constant, Roper and Scott (2009) show that UK women are more likely to

[5] See Jennings and Brush (2013) for a more comprehensive review of the literature on the entrepreneurial gender gaps and its determinants.

perceive financial barriers to start businesses than men. Self-employed women also report to have a significantly limited access to finance and being subjected to higher charges for loans in various countries (Bellucci *et al.*, 2010; De Bruin *et al.*, 2007; Kim, 2006; Marlow and Patton, 2005; Muravyev *et al.*, 2009; OECD, 2012; Guzman and Kacperczyk, 2019). Hence, the unequal access to the credit market may force women to either give up their dream of owning a business or become self-employed in occupations that are less capital-intensive altogether, as noticed by Georgellis and Wall (2000, 2005) and Campbell and Daly (1992).[6]

In this article, we investigate the extent to which a positive shock in capital has the potential to causally reduce the entrepreneurial gender gap. More specifically, we estimate whether easing the access to the capital market for women would increase their likelihood of becoming self-employed.

To test whether an increase in capital has a causal effect on the likelihood to enter self-employment, we use data from the British Household Panel Survey (BHPS) between 1997 and 2008 and investigate whether lottery wins affect individual's propensity to become self-employed. Lottery wins offer a setting that is as close as possible to a natural experiment: conditional on winning, the amount of lottery win is assumed to be randomly distributed across lottery winners. Hence, the quasi-experimental nature of lottery win allows us to test whether an increase in capital has the same causal effect in terms of direction, magnitude, and statistical significance on the probability of becoming self-

[6] Lombard (2001) also finds that although time flexibility and nonstandard work week are important, women's self-employment and the rising importance of female self-employment is mostly explained by the increase in women's earning potential as self-employed.

employment across gender (Hurst and Lusardi, 2004). While there are a few notable studies that have used lottery wins to investigate the effect of capital increase on self-employment, e.g., Lindh and Ohlsson (1996) in Sweden and Taylor *et al.* (2002) in the UK, none, to the best of our knowledge, have comprehensively investigated whether and by how much an exogenous increase in capital for women would help reduce the entrepreneurial gender gap.

Our results show that, conditional on winning, the probability of being self-employed increases with the amount of lottery win one year before. More importantly, we find no difference between men and women: the top 25% winners (an average win = £831.16) in year *t*-1 report a significant increase in the probability of self-employment in year *t* by around approximately 2 percentage points, which is 20-30% of the gender entrepreneurial gap. In contrast with previous work (e.g., Georgellis and Wall, 2005), these results thus suggest that capital constraints do play a role in explaining female self-employment and policies aimed at improving women's access to the capital market would help reduce the entrepreneurial gender gap.

The remainder of the paper is organised as follows. Section 2 discusses the conceptual framework and the previous literature. Section 3 describes the data, the empirical strategy and the estimation sample. The main results then appear in Section 4, followed by the heterogeneity analyses and robustness checks. Last, Section 5 concludes.

## 2. Lottery wins, capital constraints and gender gap in entrepreneurship

The lack of capital and limited access to the credit market are usually considered as one of the main barriers to entrepreneurship (Blanchflower and Oswald, 1998; Evans and Leighton, 1989; Evans and Jovanovic, 1989; Cagetti and De Nardi, 2006; Lindh and Ohlsson, 1996). As in Hurst and Lusardi (2004), we consider lottery wins as a source of resources that may relax capital constraints: aspiring entrepreneurs with high lottery wins would be more likely to start a business without having to look for external funding. Note that this is also true for individuals who are already self-employed and might be looking for cash flows to maintain their activity. We consider lottery wins to be an *arguably exogenous* source of new financial resources to the extent that each player has the same chance of winning.[7]

If the effect of lottery wins on the self-employment probability turn out to be null for women, which would be in line with Georgellis and Wall (2005), it would suggest that women are more likely to select themselves into less capital-intensive businesses and the gender gap in entrepreneurship would solely be caused by non-pecuniary issues. One could make a similar argument if the effect of lottery wins on the self-employment probability is positive for both men and women but significantly smaller in magnitudes for the later; under the assumption that there are no differences in access to capital, the gender gap in entrepreneurship would only reflect differences in preferences across gender.

[7] One may argue that risk-loving individuals likely spend more in lottery tickets. Although we do not observe the money spent in lottery in our empirical analysis, we show in the robustness checks that attitudes towards risk do not drive our results.

However, an extensive literature (see Roper and Scott, 2009; Bellucci *et al.*, 2010; De Bruin *et al.*, 2007; Kim, 2006; Marlow and Patton, 2005; Muravyev *et al.*, 2009; OECD, 2012; Guzman and Kacperczyk, 2019) shows that access to capital varies substantially across gender. Other things being equal, it is typically more difficult for women to get an entrepreneurial project externally funded. This means that for an entrepreneurial project of equal quality, men are more likely to get the necessary capital to start his business as compared to women. Consequently, if the effect of lottery gains on the self-employment probability is the same across gender, increasing women's access to capital should, in principle, reduce the gender gap in entrepreneurship.

## 3. Data and Empirical Strategy

### 3.1. Data

The main data source used in the analysis is the British Household Panel Survey (BHPS).[8] This is a nationally representative random sample of over 10,000 household, containing almost 30,000 unique adult individuals, conducted between September and Christmas of each year from 1991 to 2008 (see Taylor *et al.*, 2002). Data on self-employment have been collected from 1991, whilst data on lottery wins were collected for the first time in 1997.

The exact question asked about lottery win is, "*About how much in total did you receive? Win on the football pools, national lottery or other form of gambling?*". In modern Britain, the national lottery is overwhelmingly the main form of gambling

[8] The BHPS is freely available for download after registration at www.data-archive.ac.uk.

relevant to this question, so for succinctness we shall refer to this as lottery wins. For the design of this study, any form of gambling windfall would be suitable as a quasi-experimental income shock since the amount received is randomly distributed across winners.

Around 36% of the BHPS working age adults, i.e., 16-65 years old, report at least one lottery win. After adjusting for prices using consumer price index (CPI, year=2000), 81% of these lottery wins are small wins (£1-£99), 14% are medium-sized wins (£100-£499), and 5% are big wins (£500+). The average win for the top 25% winners is £831.16 (minimum = £85.66; maximum = £184,672.20), and the average win for the bottom 75% is £25.69 (minimum = £1.64; maximum = £85.23). Many people also won money from the lottery more than once, with an average of around 2 wins per person who reported to have won lottery in the panel. However, the BHPS does not contain information about the number of times (if any) the individual has played the lottery. Hence, we cannot distinguish non-players from unsuccessful players.

We define self-employed as being those who declare being self-employed in survey year *t*. Among our potential pool of self-employed, we include all individuals who are in working age, i.e., 16-65 years old (N=189,820 observations; n=28,042 unique individuals). The share of self-employed in our entire working-age sample is 8.1%.[9] Self-employed have won lottery as much as the rest of the population, that is 36% of them report at least one lottery win. However, conditional on winning, they have won bigger

[9] According to the ONS, the share of self-employed in 2001 in the UK was 12%. However, the definition of the labor force population used by the ONS to produce this figure is more restrictive.

amounts on average. The average, inflation-adjusted winning for self-employed is £595.50 (S.D.=5,541.33) with the maximum win of £169,635, while for the rest of the sample is £192.79 (S.D.=£2549.44) and the maximum win of £184,672.

We can also examine whether there is any significant difference by gender. Figure 1 illustrates the distribution of lottery wins by gender and self-employment status. The average winning for women who are self-employed is £197.57 (S.D.=621.59) while the corresponding figure among the rest of the female sample is £168.91 (S.D.=2,313.59). This is respectively £696.92 (S.D.=6,196.11) and £211.99 (S.D.=2,724.23) for men.

### 3.2. Empirical strategy

Our main self-employment regression equation can be written as follows:

$$S_{it} = \alpha + \beta_1 FEMALE_i + \beta_2 Top25\%_{it} + \beta_3 (FEMALE_i \times Top25\%_{it}) + X'_{it}\gamma + u_i + \varepsilon_{it} \quad (1)$$

where $S_{it}$ denotes self-employment status (e.g., an indicator variable with 0 = not self-employed and 1 = self-employed) for individual *i* aged 16-65 years old, in survey year *t*; $Top25\%_{it}$ is a dummy representing the top 25% of winner in the lottery, which corresponds to an average win of £831.16 (minimum = £85.66; maximum = £184,672.20), for individual *i* in year t (takes the value 0 if individual *i* won a lottery but not in the amount that would qualify her in the top 25% of lottery income distribution,

i.e., average win = £25.69 (minimum = £1.64; maximum = £85.23))[10]; $FEMALE_i$ is an indicator variable with 0 = male and 1 = female; $X'_{it}$ is a vector of personal and household characteristics, including age, age-squared, log of real equivalent income, marital status, highest completed education level, self-assessed health status, number of days spent in hospital last year, number of dependent children, home ownership, regional dummies, and time dummies; $u_i$ denotes individual-specific effects, and $\varepsilon_{it}$ is the error term.

Eq.1 sets out to test the following three main hypotheses. First, consistent with previous studies that find a higher incidence of self-employment among men than among women (e.g., Georgellis and Wall, 2005; Koellinger *et al.*, 2013) we expect that $\beta_1 < 0$. Second, consistent with the liquidity constraint hypothesis (e.g., Lindh and Ohlsson, 1996; Blanchflower and Oswald, 1998; Taylor *et al.*, 2002), we anticipate that $\beta_2 > 0$. Finally, consistent with the assumption that capital has the same marginal effect on both male and female aspiring to become entrepreneurs, we hypothesise that $\beta_3 = 0$.

Given that it may take some time for people to make a transition from other job statuses to self-employment following a lottery win, we also rewrite Eq.1 to include a lag lottery win as the main explanatory variable as follows:

$$S_{it} = \alpha + \beta_1 FEMALE_i + \beta_2 Top25\%_{it-1} + \beta_3 (FEMALE_i \times Top25\%_{it-1}) + X'_{it}\gamma + u_i + \varepsilon_{it} \quad (2)$$

[10] Given that the majority of the win is under £100, we follow one of the referees' advice and focus our attention mainly on the large winners, i.e., the top 25% winners.

All regressions, unless stated, are estimated using random effects logit model with standard errors clustered at the individual level. Given that logit coefficients are not directly interpretable as marginal effects, we also calculate and report the relevant marginal effects at the mean in all tables. For robustness checks, we also perform linear regression models, which are reported in the Appendix.

Our first estimation sample is made of the working age adults, i.e., 16-65 years old, with valid information for the dependent, independent and control variables. In these regressions, $\beta_2$ and $\beta_3$ are causal under the assumption that lottery gains are randomly distributed across the working-age population. However, we can see from the summary statistics reported in Table 1A in the Appendix that lottery winners have, on average, statistically significantly different characteristics from the population of non-winners. For example, both female and male lottery winners at the year of winning are likely to be older, have higher household incomes, have worse self-assessed health, and own home outright than their non-winner counterparts.

To minimise any bias that might arise from the systematic differences between lottery players and non-players, we further restrict our sample to include only lottery winners *at the year of their winning* to estimate our baseline estimates. This produces 13,934 observations (6,097 individuals). Of those, 5,929 observations (2,765 individuals) are women, and 1,162 observations (569 individuals) are self-employed. Hence, our identification strategy relies on the assumption that the amount of lottery win at the year of winning is randomly distributed across winners. Table 2A in the Appendix provides evidence on the exogeneity of these lottery gains by showing the results of regressions of

the amount of lottery win on lagged individual and household characteristics. The results show that, overall lagged individual and household characteristics are not significantly correlated with lottery gains in 16 out of 22 characteristics tested, thus confirming that the amount of lottery win is randomly distributed across winners.

In addition to Table 2A, we also illustrate in Figure 2 the dynamics of self-employment before, during, and after a large win (top 25% winners) compared to a small win (bottom 75% winners). Here, each observational point on Figure 2 represents the estimated linear probability coefficient of self-employment in the years before winning (t-3 to t-1), the year of winning (t), and the years following the win (t+1 to t+3). We can see that the probability of self-employment is mostly statistically insignificantly different from zero in the years before winning big in the lottery compared to winning small, i.e., there is no lead effects to becoming self-employed in the years before winning in the lottery. However, the probability of self-employment is positive and statistically significantly different from zero at the 5% level from year t+1 to t+3, thus reaffirming the exogeneity of lottery win in our data.

Finally, we specifically test whether lottery wins in *t-1* affect transitions into self-employment. To do so, we reproduce Eq. 2, but excluding lottery winners who were already self-employed in *t-1*. As such, our dependent variable reflects solely the probability of becoming self-employed among lottery winners.

## 4. Results

### 4.1. Main results

We begin our analysis by examining whether there is a significant gender entrepreneurial gap in our British raw data. Using the entire working-age sample (N=189,820 observations; n=28,042 unique individuals), Table A1 in Appendix shows that only 4% of the female sample compared to 12.7% of the male sample are self-employed. Approximately the same proportions of self-employed can be found among female (3.9%) and male (11.5%) lottery winners (N=13,988 observations; n=6,116 unique individuals). These aggregate numbers suggest that men are substantially more likely than women to be self-employed, which is consistent with previous findings in the literature (e.g., Georgellis and Wall, 2005; Koellinger *et al.*, 2013).

We formally estimate the extent of the gender entrepreneurial gap and assess whether lottery wins can reduce it by estimating Eqs 1 and 2. Results are reported in Table 1.[11] The first two columns use all working age individuals taken from the BHPS. The rest of the columns use only the sample of lottery winners at the year of winning. We also report the estimated marginal effects at the end of each columns.

Using the entire working age sample, Column 1 of Table 1 shows that women are *ceteris paribus* 7.2 percentage points less likely to be self-employed than men, a difference which is also statistically significant at the 1% level. We then ask in Column 2 whether a large win (i.e., top 25% winner of all winners with an average win = £831.16) compared to the small win and zero win combined influences the probability of being

[11] Full results are reported in Table 3A in appendix.

self-employed in the same estimation sample. As revealed by the figure in the bottom Panel, a large win in *t* significantly increases the probability of being self-employed in the same period by 0.6 percentage point. This positive and significant estimate is consistent with the idea that we could significantly increase the number of entrepreneurs by simply improving capital access.

Although we control for a number of covariates to attenuate concerns of endogeneity, we cannot rule out the fact that the effect of the lottery wins in Column 2 is also capturing the influence of unobserved differences between players and non-players. Hence, we re-estimate in Column 3 the same regression equation using only lottery winners in the panel. Conditional on being a winner, the amount won at the lottery is arguably exogenous and the estimate of the real lottery wins can be considered as the causal effect of income on self-employment. Results in Column 3 turn out to be similar to the previous estimates: roughly the same gender entrepreneurial gap, i.e., 6.1 percentage points, is observed among male and female lottery winners. Lottery win in period *t*, which is the same period as the self-employment status, is positive albeit statistically insignificantly different from zero.

Column 4 of Table 1 tests whether the effect of lottery win on self-employment is the same for men as for women. Looking at the interaction term between the female dummy variable and the dummy representing top 25% winners in period *t*, we find little evidence that the marginal effect of lottery win on self-employment is statistically significantly different across gender. Although the marginal effects are not significant at

conventional levels, there is little evidence to suggest that a large lottery win has a different effect on the propensity to become self-employed between men and women.

Columns 5 and 6 of Table 1 replaces top 25% winners dummy with its first lag, which allows for the possibility that it might take some time for people to become self-employed following a positive shock in capital. Both columns' results suggest that there is a statistically important lag effect of lottery win on self-employment that is the same for women as for men; being a top 25% winner in period *t*-1 increases the likelihood of self-employment in period *t* by approximately 2.2 percentage points for men and 1.4 percentage points for women. For women, this is roughly 20% of the gender entrepreneurial gap. These are economically sizeable estimates when compared to the estimated effect of homeownership on the probability of self-employment of 1.8 percentage points or the effect of an additional children on self-employment of 1 percentage point.

In a similar vein, we ask whether lottery wins in *t-1* also affect transitions into self-employment. To do so, we reproduce our main regressions but excluding individuals who were already self-employed in *t-1*. As such, our dependent variable reflects solely the probability of becoming self-employed. Table 2 displays the results. Similar results are obtained, although the magnitudes are now slightly smaller.

While previous works suggest that only men are constrained by the lack of capital endowment (e.g., Georgellis and Wall, 2005), evidence of similar marginal effects in Table 1 and 2 across gender suggests that women gain as much as men from a positive

capital shock, which also implies that women may also face a similar liquidity concern as much as men in their decision to become self-employed when they lack capital.

### 4.2. Heterogeneity analysis

Table 3 investigates whether the gender entrepreneurial gap and the effects of lottery gains are larger for certain types of lottery winners by re-estimating Eq.2 on different subsamples. More specifically, we look for differences across age, education, household income, marital status, parenthood and homeownership. Overall, we continue to find evidence of a significant gender entrepreneurial gap in eleven out of twelve subsamples. Men are significantly more likely than women to be self-employed across age groups, across education and income groups, married and cohabiting individuals, both parents and non-parents, and homeowners and renters.

We also show that a one-year lag lottery win has a positive and statistically significant effect on the likelihood of becoming self-employed across almost all subsamples for men. Only men with a rent are not more likely to become self-employed after winning at the lottery. This is not surprising as windfall gains might be more likely to be invested in real estate first.

There is more heterogeneity when we look at the marginal effects for female. On average, we find that lottery wins increase the probability of becoming self-employed for women who are aged 40 and above, with low levels of education and income (below the median), with a partner or a husband, with children and who own their house. This could suggest that for these women, self-employment can be seen as a substitute for part-time

work or labour market inactivity, allowing them to have time-flexibility and a greater opportunity to deal with household/childcare responsibilities.

One of the largest gender entrepreneurial gaps at approximately 9 percentage points is found amongst the older age group, i.e., age 40-65 years old. For this subgroup, being a top 25% winner in period *t*-1 increases the likelihood of becoming self-employed in period *t* by approximately 2 percentage points for females, which implies that the average lottery win has to be around four times higher for these women to fully close the gender entrepreneurial gap.

Table 4 addresses the question of whether being married or cohabiting with someone who wins lottery matters. For example, would a husband's win in period *t-1* increases the probability that his wife will be self-employed in period *t*, as would have been predicted by the household collective models (e.g., Chiappori, 1992). However, looking at Table 4's results, we find little evidence that a spouse's win in period *t-1* increases the probability of self-employment in period *t* for the other spouse. Hence, we have evidence that the effect of lottery win on self-employment is only present among the winners and not the winner's partner or spouse.

### 4.3. Robustness checks

We carry out several robustness checks and report them in the appendix. First, we discuss the sensitivity of our results with respect to the choice of the estimation method. To do so, we replicate our analysis using a linear probability model instead of the logit with random effects. Results are shown in Table 4A in the appendix. In line with the baseline

estimates, we find a significant gender entrepreneurial gap of roughly 7 percentage points and a positive and significant estimate for lottery wins that is the same for men and women.

Second, a causal interpretation of our main estimates assumes that self-employment and lagged lottery wins are not simultaneously determined by omitted variables. Given that most entrepreneurs tend to have higher propensity for risks than an average person (see, e.g., Stewart and Roth, 2001), one may argue that they might buy more lottery tickets and, consequently, have a higher likelihood of winning and reporting higher wins. While the amount of money spent in lottery tickets is not reported in BHPS, respondents were asked to reply to the following question: "*Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?*" using a Likert scale ranging from 1 ("*Won't take risks*") to 10 ("*Ready to take risks*"). We consider the response to be a reasonable proxy for the general risk-taking attitude, which should correlate with how much individuals typically spend on lottery tickets. Because this risk attitude variable was asked in the last wave of BHPS only, we use it to create a time-invariant individual-specific risk attitude variable that we then use as a control variable in our self-employment regression. We first reproduce in the first column of Table 5A our main estimates for the individuals with non-missing risk attitude. Results in this subsample (25% smaller than the main estimation sample) are the same: being a top 25% winner in t-1 increases the probability of self-employment for both men and women by approximately 2 percentage points. This result implies that the relationship between

lottery gains and the probability of being self-employed we observed in Table 1 is not spuriously driven by the influence of unobserved differences in risk aversion.

Third, given that price-adjusted lottery wins in our estimation sample range from £0.82 to £184,672, we next investigate whether more categories of "small" v.s. "large" gains have a different impact on the probability to be self-employed. Rather than defining "small" and "large" gains arbitrary on the basis of arbitrary thresholds, we generate four categories of lottery wins based on the quartiles of the lottery gains distribution coming from our estimation sample (Q1=£1-£26.3; Q2=£26.3-£77.2; Q3=£77.2-£230.6; Q4= above £230.6). Table 6A shows the results. In the first column, we present the estimate on log real lottery win in year t-1, which shows that lottery win in its log form is positively and statistically significantly correlated with the probability of self-employment in year t. In column (2), we show the results when we replace the continuous lottery wins (in log) by categories of lottery gains. We find again that none of the interaction terms attracts a significant estimate: the effects of lottery gains are the same across gender. Nevertheless, it appears that only wins in the top quartile significantly predict a larger self-employment probability. This is not surprising: small lottery wins are unlikely to be sufficient to relax capital constraints.

It might also be more appropriate to replace the contemporaneous covariates by their one-year lag as it allows us to control for the lagged employment status and avoid "bad controls" as discussed in Angrist and Pischke (2008). We demonstrate in Table 7A doing so produces similar coefficients, although the standard errors are now somewhat larger with lagged control variables.

Rather than using interaction terms to account for gender differences, we could use sample-splits. Table 8A, which shows the results separately for men and women, demonstrates that lottery win in period *t-1* produces similar coefficients on the likelihood of self-employment in period *t* across gender.

Finally, Table 9A examines the longer lag effects of lottery win and finds that a win in period *t-2* has a positive effect on self-employment in period t for both men and women, although the effect is non-significant at conventional levels. On the other hand, the effect of lottery win in period *t-3* is only positive and statistically significant for men and not women.

## 5. Conclusions

Using the BHPS, we show that lottery gains, an exogenous source of resources, predict a significant increase in the probability of becoming self-employed. Seen through the lenses of the theory of capital constraints on potential entrepreneurs, this result is not surprising: lottery gains are new resources that can be used to compensate insufficient capital endowment. We also find that the effect of lottery gains is the same for men and women on average. In contrast with previous work (e.g., Georgellis and Wall, 2005), this suggests that capital constraints matter as much to women as men in determining the decision of whether to enter self-employment. Finally, the effect of lottery gains is roughly the same for all types of men while it is only significant for women who have less resources, partnered and with children at home.

We believe that these results have important policy implications. For instance, our findings suggest that the gender entrepreneurial gap is not only due to differences in preferences across gender (e.g., women being more risk-averse), but might also reflect a lack of sufficient access to the capital market for women. Consequently, policies aimed at easing the access to the capital market for women might help to close the gender entrepreneurial gap. Many countries have introduced government programmes aimed at providing transfer payments to individuals who want to become entrepreneurs. Our results should provide some new insights into whether those policies will be efficient in increasing the number of women who choose to become entrepreneur.

Like all studies in social sciences, our study is not without limitations. Although lottery gains are arguably exogenous among winners, lottery players are likely to have unobservable characteristics significantly different from the rest of the population. While this is not problematic regarding the internal validity of our analysis, its external validity might be somehow limited as only slightly more than half of the British adult population plays the lottery. Future research will have to return to investigate the effect of other types of positive income shocks, one that affect the population more generally, on each gender's propensity to enter self-employment.

**References:**


Angrist, J. D., and Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Ardagna, S., and Lusardi, A. (2008). *Explaining international differences in entrepreneurship: The role of individual characteristics and regulatory constraints*. National Bureau of Economic Research Working Paper Series, No. 14012.

Bellucci, A., Borisov, A., and Zazzaro, A. (2010). "Does gender matter in bank–firm relationships? Evidence from small business lending", *Journal of Banking and Finance, 34*, 2968-2984.

Blanchflower, D. G. (2000) "Self-employment in OECD countries", *Labour Economics, 7*, 471-505.

Blanchflower, D. G. and Meyer, B. D. (1994) "A longitudinal analysis of the young self-employed in Australia and the United States", *Small Business Economics, 6,* 1-19.

Blanchflower, D. G., and Oswald, A. J. (1998). "What makes an entrepreneur?", *Journal of Labor Economics, 16,* 26-60.

Cagetti, M., and De Nardi, M. (2006). "Entrepreneurship, frictions, and wealth", *Journal of Political Economy, 114*, 835-870.

Caliendo, M., Fossen, F., and Kritikos, A. S. (2014). "Personality characteristics and the decisions to become and stay self-employed", *Small Business Economics, 42*, 787-814.

Campbell, M. and Daly, M. (1992) "Self-employment into the 1990s", *Employment Gazette,* 269-292.

Chiappori, P. A. (1992). “Collective labor supply and welfare”, *Journal of Political Economy*, *100*, 437-467.

Clark, A.E., and Lepinteur, A. (2020). “Making the grades? The cognitive and non-cognitive antecedents of self-employment”, *Mimeo*

Colombier, N., and Masclet, D. (2008). “Intergenerational correlation in self-employment: Some further evidence from ECHP Data”, *Small Business Economics,* 30, 423-437.

Connelly, R. (1992). “The effect of child care costs on married women's labor force participation”, *Review of Economics and Statistics, 74,* 83-90.

Cramer, J. S., Hartog, J., Jonker, N., and Van Praag, C. M. (2002). “Low risk aversion encourages the choice for entrepreneurship: an empirical test of a truism”, *Journal of Economic Behavior and Organization, 48,* 29-36.

De Bruin, A., Brush, C. G., and Welter, F. (2007). “Advancing a framework for coherent research on women's entrepreneurship”, *Entrepreneurship Theory and Practice, 31,* 323-339.

Djankov, S., Qian, Y., Roland, G., and Zhuravskaya, E. (2008). “What Makes an Entrepreneur?”, *International Differences in Entrepreneurship*.

Evans, D. S., and Jovanovic, B. (1989). “An estimated model of entrepreneurial choice under liquidity constraints”, *Journal of Political Economy, 97,* 808-827.

Evans, D. S., and Leighton, L. S. (1989). “The determinants of changes in US self-employment, 1968–1987”, *Small Business Economics, 1,* 111-119.

Fairlie, R. W., and Robb, A. (2007). “Families, human capital, and small business: Evidence from the characteristics of business owners survey”, *ILR Review, 60*, 225-245.

Fujii, E. T., and Hawley, C. B. (1991). “Empirical aspects of self-employment”, *Economics Letters, 36,* 323-329.

Georgellis, Y., and Wall, H. J. (2000). “What makes a region entrepreneurial? Evidence from Britain”, *The Annals of Regional Science, 34*, 385-403.

Georgellis, Y., and Wall, H. J. (2005). “Gender differences in self-employment”, *International Review of Applied Economics, 19,* 321-342.

Guzman, J., and Kacperczyk, A. O. (2019). “Gender gap in entrepreneurship”, *Research Policy*, *48*, 1666-1680.

Hurst, E., and Lusardi, A. (2004). “Liquidity constraints, household wealth, and entrepreneurship”, *Journal of Political Economy, 112*, 319-347.

Jennings, J. E., and Brush, C. G. (2013). “Research on women entrepreneurs: challenges to (and from) the broader entrepreneurship literature?”, *Academy of Management Annals, 7,* 663-715.

Kim, G. O. (2006). “Do equally owned small businesses have equal access to credit?”, *Small Business Economics, 27,* 369-386.

Koellinger, P., Minniti, M., and Schade, C. (2013). “Gender differences in entrepreneurial propensity”, *Oxford Bulletin of Economics and Statistics*, *75*, 213-234.

Lindh, T., and Ohlsson, H. (1996). “Self-employment and windfall gains: evidence from the Swedish lottery”, *Economic Journal, 106,* 1515-1526.

Lombard, K. V. (2001). “Female self-employment and demand for flexible, nonstandard work schedules”, *Economic Inquiry, 39,* 214-237.

Marlow, S., and Patton, D. (2005). “All credit to men? Entrepreneurship, finance, and gender”, *Entrepreneurship Theory and Practice, 29,* 717-735.

Muravyev, A., Talavera, O., and Schäfer, D. (2009). “Entrepreneurs' gender and financial constraints: Evidence from international data”, *Journal of Comparative Economics, 37,* 270-286.

OECD. (2012). “Gender equality in education, employment and entrepreneurship: Final report to the MCM 2012”, *Better Policies for Better Lives*.

Rees, H., and Shah, A. (1986). “An empirical analysis of self-employment in the UK”, *Journal of Applied Econometrics, 1,* 95-108.

Roper, S., and Scott, J. M. (2009). “Perceived financial barriers and the start-up decision: An econometric analysis of gender differences using GEM data”, *International Small Business Journal, 27,* 149-171.

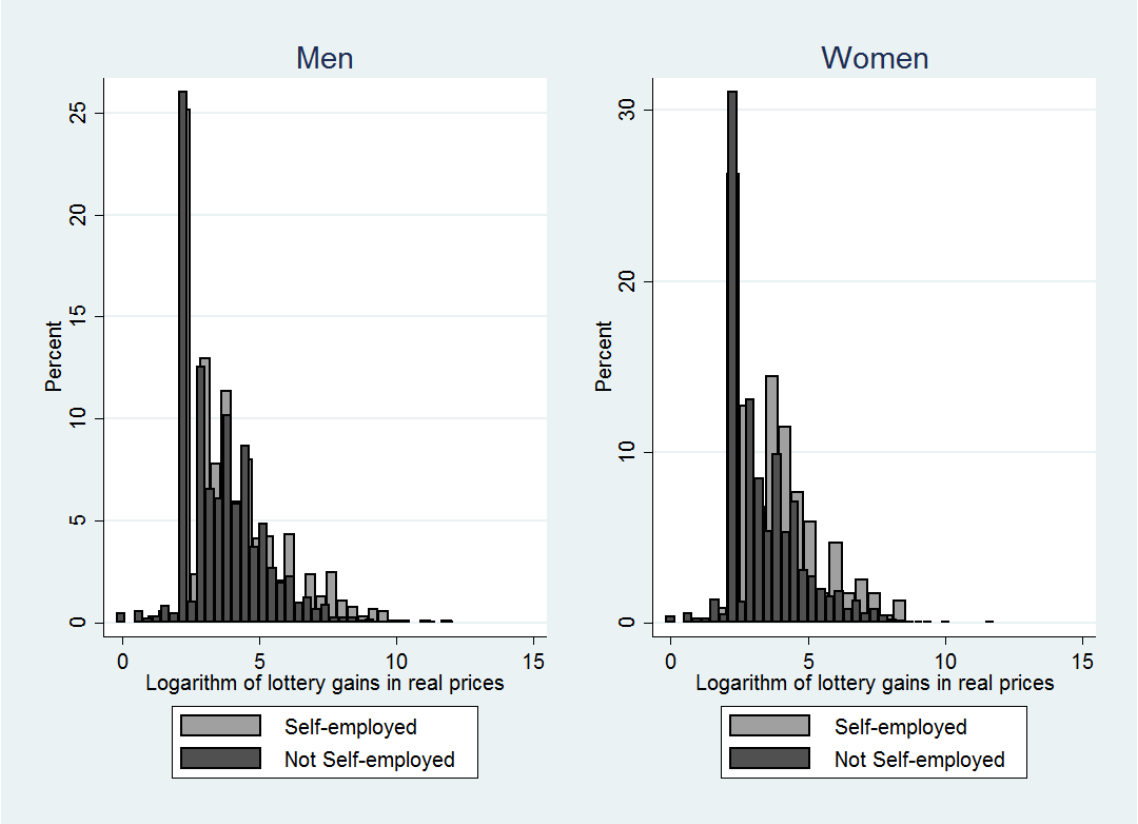
Stewart, W. H., Jr., & Roth, P. L. (2001). Risk propensity differences between entrepreneurs and managers: A meta-analytic review. *Journal of Applied Psychology, 86*(1), 145–153.

Taylor, M. P. (1999). “Survival of the fittest? An analysis of self-employment duration in Britain”, *Economic Journal, 109,* 140-155.

Taylor, M. F., Brice, J., Buck, N. and Prentice-Lane, E. (2002) *British Household Panel Survey User Manual*. Colchester: University of Essex.

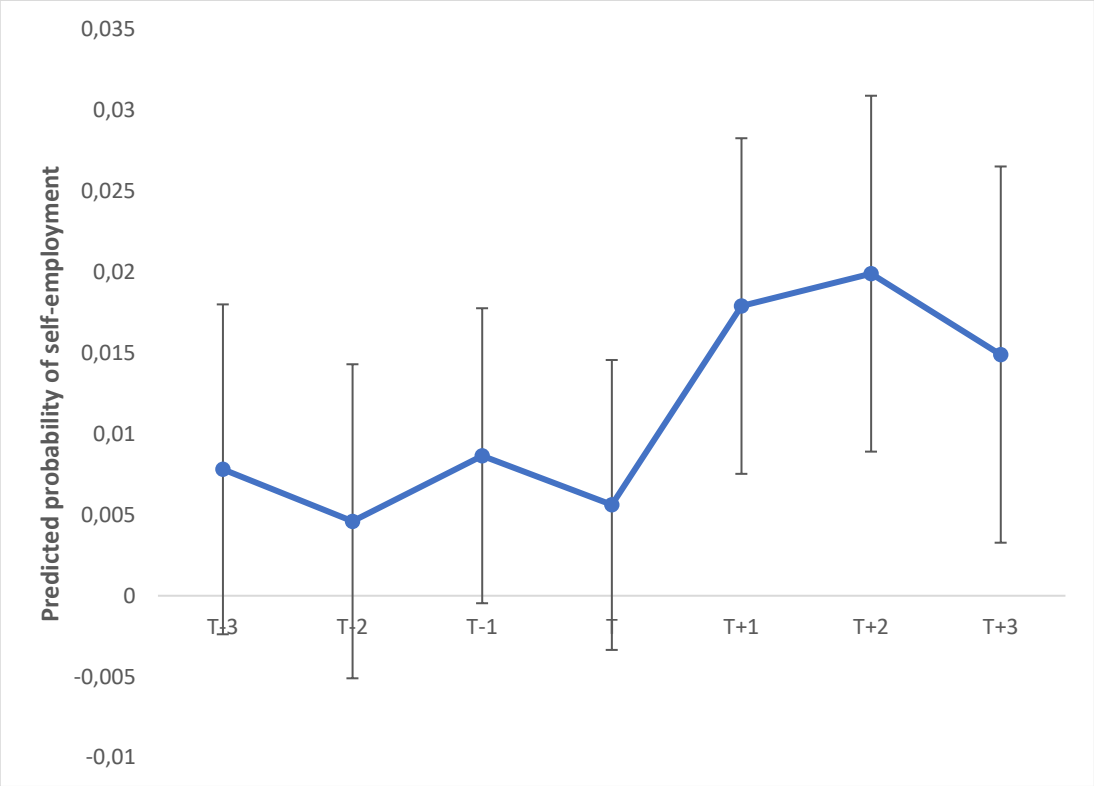
**Figure and Tables:**

Figure 1: Distribution of lottery wins by gender and self-employment status




Note: Lottery gains are in pounds adjusted using consumer price index (CPI, year=2000) and in logarithmic form.

**Figure 2: Leads and lags to self-employment for top 25% winner in the lottery: Linear probability model with random effects**




Notes: 4-standard-error bars (2 S.E. above, 2 S.E. below: 95% C.I.). Each data point represents the estimated coefficient on top 25% winner in years before winning (t-3 to t-1), the year of winning (t), and the years following the win (t+1 to t+3). The reference group for each data point is the bottom 75% winner. We ran seven separate regressions to obtain each of the seven lead and lag coefficients. This is because there are very few winners with observations from all seven time points. For simplicity, we used linear probability with random effects model to obtain these coefficients. Number of observations used in the estimation: t-3 (N=10,818), t-2 (N=11,661), t-1 (N=12,629), t (N=13,934), t+1 (N=11,952), t+2 (N=10,401), and t+3 (N=9,013).

**Table 1: Self-employment and lottery wins: Logit with random effects regressions**

| | Self-employment Probability in *t* | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Female | -2.515*** | -2.513*** | -2.518*** | -2.586*** | -2.353*** | -2.516*** |
| | (0.0778) | (0.0778) | (0.271) | (0.291) | (0.386) | (0.409) |
| Top 25% winner in period t | | 0.230** | 0.182 | 0.110 | | |
| | | (0.107) | (0.155) | (0.169) | | |
| Female × Top 25% winner in period t | | | | 0.259 | | |
| | | | | (0.389) | | |
| Top 25% winner in period t-1 | | | | | 0.661*** | 0.547*** |
| | | | | | (0.180) | (0.193) |
| Female × Top 25% winner in period t-1 | | | | | | 0.423 |
| | | | | | | (0.452) |
| **Marginal effects at the mean** | | | | | | |
| Female | -0.072*** | -0.072*** | -0.061*** | -0.061*** | -0.060*** | -0.062*** |
| | (0.002) | (0.002) | (0.006) | (0.006) | (0.007) | (0.006) |
| Top 25% winner in period t | | 0.006** | 0.005 | | | |
| | | (0.003) | (0.004) | | | |
| Male: Top 25% winner in period t | | | | 0.004 | | |
| | | | | (0.006) | | |
| Female: Top 25% winner in period t | | | | 0.004 | | |
| | | | | (0.004) | | |
| Top 25% winner in period t-1 | | | | | 0.018*** | |
| | | | | | (0.005) | |
| Male: Top 25% winner in period t-1 | | | | | | 0.022*** |
| | | | | | | (0.008) |
| Female: Top 25% winner in period t-1 | | | | | | 0.014** |
| | | | | | | (0.007) |
| Control variables | Yes | Yes | Yes | Yes | Yes | Yes |
| Regional dummies | Yes | Yes | Yes | Yes | Yes | Yes |
| Wave dummies | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 189,820 | 189,820 | 13,934 | 13,934 | 11,952 | 11,952 |
| Number of individuals | 28,042 | 28,042 | 6,097 | 6,097 | 5,331 | 5,331 |

Notes: The dependent variable is a dummy variable that takes the value one if the respondent is self-employed in period *t* and zero otherwise. The figures in Columns 1 and 2 refer to the working-age respondents while figures in the remaining Columns refer to the lottery winners in the year of winning. The average real winning for the top 25% winners = £831.16 (minimum = £85.66; maximum = £184,672.20), and the bottom 75% winners = £25.69 (minimum = £1.64; maximum = £85.23). Control variables in each regression are the age, the age squared, the log of real equivalent household income, dummies for marital status, dummies for self-reported health status, education dummies, homeownership, the number of days spent in hospital in year t-1 and the number of dependent children. Standard errors are clustered at the individual level. *<10%; **<5%; ***<1%.

**Table 2: Transition into self-employment and lottery wins in t-1: Logit with random effects regressions**

| | Self-employment Probability in *t*, conditioning on not being in Self-employment in *t*-1 | |
|---|---|---|
| | (1) | (2) |
| Female | -0.755*** | -0.744*** |
| | (0.224) | (0.256) |
| Top 25% winner in period t-1 | 0.677*** | 0.688*** |
| | (0.194) | (0.234) |
| Female × Top 25% winner in period t-1 | | -0.0361 |
| | | (0.411) |
| **Marginal effects at the mean** | | |
| Female | -0.010*** | -0.010*** |
| | (0.003) | (0.003) |
| Top 25% winner in period t-1 | 0.009*** | |
| | (0.003) | |
| Male: Top 25% winner in period t-1 | | 0.013*** |
| | | (0.005) |
| Female: Top 25% winner in period t-1 | | 0.007* |
| | | (0.004) |
| Control variables | Yes | Yes |
| Regional dummies | Yes | Yes |
| Wave dummies | Yes | Yes |
| Observations | 10,620 | 10,620 |
| Number of individuals | 4,890 | 4,890 |

Notes: The dependent variable is a dummy variable that takes the value one if the respondent is self-employed in period *t* and zero otherwise. The figures refer to the lottery winners a year after of winning who were not self-employed in *t-1*. The average real winning for the top 25% winners = £831.16 (minimum = £85.66; maximum = £184,672.20), and the bottom 75% winners = £25.69 (minimum = £1.64; maximum = £85.23). Control variables in each regression are the age, the age squared, the log of real equivalent household income, dummies for marital status, dummies for self-reported health status, education dummies, homeownership, the number of days spent in hospital in year *t-1* and the number of dependent children. The regressions also include region and wave dummies. Standard errors are clustered at the individual level. *<10%; **<5%; ***<1%.

**Table 3: Self-employment and lottery wins across different subsamples: Logit with random effects regressions**

| | Self-employment Probability in *t* | | | | | |
|---|---|---|---|---|---|---|
| | *Young* | *Old* | *High education* | *Low education* | *High income* | *Low income* |
| Female | -2.097*** | -2.918*** | -2.432*** | -2.716*** | -2.633** | -2.513*** |
| | (0.503) | (0.493) | (0.526) | (0.476) | (1.101) | (0.359) |
| Top 25% winner in period t-1 | 0.635* | 0.595*** | 0.561* | 0.572** | 0.211 | 0.727*** |
| | (0.330) | (0.224) | (0.289) | (0.256) | (0.313) | (0.261) |
| Female × Top 25% winner in period t-1 | 0.058 | 0.633 | 0.650 | 0.231 | 0.353 | 0.246 |
| | (0.678) | (0.545) | (0.607) | (0.629) | (0.845) | (0.529) |
| **Marginal effects at the mean** | | | | | | |
| Female | -0.036*** | -0.090*** | -0.062*** | -0.062*** | -0.062*** | -0.069*** |
| | (0.006) | (0.0092) | (0.0082) | (0.0107) | (0.019) | (0.0088) |
| Male: Top 25% winner in period t-1 | 0.017* | 0.022** | 0.022* | 0.023** | 0.008 | 0.032*** |
| | (0.010) | (0.009) | (0.011) | (0.010) | (0.012) | (0.012) |
| Female: Top 25% winner in period t-1 | 0.007 | 0.027** | 0.021* | 0.010 | 0.008 | 0.015* |
| | (0.007) | (0.013) | (0.012) | (0.008) | (0.011) | (0.008) |
| Observations | 5,736 | 6,216 | 4,681 | 7,148 | 5,978 | 5,974 |
| Number of individuals | 2,889 | 2,786 | 2,220 | 3,114 | 2,964 | 3,284 |
| | Self-employment Probability in *t* | | | | | |
| | *Partnered* | *Not partnered* | *No children* | *With children* | *Renters* | *Homeowners* |
| Female | -2.478*** | -2.279*** | -2.663*** | -2.564*** | -4.035*** | -2.366*** |
| | (0.335) | (0.714) | (0.725) | (0.599) | (1.401) | (0.363) |
| Log (real lottery win) in period t-1 | 0.517** | 1.220*** | 0.801*** | 0.435 | 0.948* | 0.534*** |
| | (0.202) | (0.445) | (0.240) | (0.355) | (0.547) | (0.201) |
| Female× Log (real lottery win) in period t-1 | 0.471 | -0.560 | -0.0136 | 1.145 | -0.636 | 0.478 |
| | (0.434) | (1.218) | (0.643) | (0.697) | (1.435) | (0.454) |
| **Marginal effects at the mean** | | | | | | |
| Female | -0.076*** | -0.031*** | -0.054*** | -0.061*** | -0.033** | -0.065*** |
| | (0.007) | (0.008) | (0.008) | (0.009) | (0.012) | (0.007) |
| Male: Top 25% winner in period t-1 | 0.021** | 0.030** | 0.030*** | 0.014 | 0.021 | 0.022*** |
| | (0.008) | (0.013) | (0.009) | (0.012) | (0.015) | (0.008) |
| Female: Top 25% winner in period t-1 | 0.0208** | 0.006 | 0.008 | 0.016** | 0.002 | 0.019** |
| | (0.009) | (0.011) | (0.006) | (0.008) | (0.010) | (0.009) |
| Observations | 8,750 | 3,202 | 8,143 | 3,809 | 2,565 | 9,387 |
| Number of individuals | 3,894 | 1,716 | 3,834 | 1,884 | 1,437 | 4,142 |

Notes: The dependent variable is a dummy variable that takes the value one if the respondent is self-employed in period *t* and zero otherwise. The figures refer to the working-age lottery winners the year of winning. "Young" and "Old" are, respectively, respondents below age 40 and between 40 and 65 years old. "High Education" and "Low Education" are respectively respondents with and without a university degree. "Low Income" and "High Income" are respectively respondents with a log of real equivalent household income below and above the median of the log of real equivalent household income of the estimation sample. The figures refer to the lottery winners a year after of winning who were not self-employed in *t-1*. The average real winning for the top 25% winners = £831.16 (minimum = £85.66; maximum = £184,672.20), and the bottom 75% winners = £25.69 (minimum = £1.64; maximum = £85.23). Control variables in each regression are the age, the age squared, the log of real equivalent household income, dummies for marital status, dummies for self-reported health status, education dummies, homeownership, the number of days spent in hospital in year t-1 and the number of dependent children. The regressions also include region and wave dummies. Standard errors are clustered at the individual level. *<10%; **<5%; ***<1%.

**Table 4: Self-employment and spouse's lottery win: Logit with random effects regressions**

| | Self-employment Probability in *t* |
|---|---|
| Female | -2.618*** |
| | (0.851) |
| Spouse: Top 25% winner in period t-1 | 0.0513 |
| | (0.261) |
| Female × Spouse: Top 25% winner in period t-1 | -0.192 |
| | (0.472) |
| **Marginal effects at the mean** | |
| Female | -0.077*** |
| | (0.007) |
| Female: Spouse: Top 25% winner in period t-1 | 0.002 |
| | (0.009) |
| Male: Spouse: Top 25% winner in period t-1 | -0.002 |
| | (0.005) |
| Observations | 8,186 |
| Number of individuals | 3,629 |

Notes: The dependent variable is a dummy variable that takes the value one if the respondent is self-employed in period *t* and zero otherwise. The figures refer to the working-age respondents partnered to lottery winners the year of winning. Control variables in each regression are the age, the age squared, the log of real equivalent household income, dummies for marital status, dummies for self-reported health status, education dummies, homeownership, the number of days spent in hospital in year *t-1* and the number of dependent children. Standard errors are clustered at the individual level. The regressions also include region and wave dummies. *<10%; **<5%; ***<1%.

## Appendix

### Table 1A: Summary statistics

| | Women | | | | | Men | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Non-winners | | Winners | | | Non-winners | | Winners | | |
| | M | S.E. (M) | M | S.E. (M) | t-test | M | S.E. (M) | M | S.E. (M) | t-test |
| Self-employed | 0.040 | 0.001 | 0.039 | 0.003 | | 0.128 | 0.001 | 0.115 | 0.004 | *** |
| Age | 39.154 | 0.044 | 41.491 | 0.172 | *** | 39.082 | 0.048 | 39.46 | 0.149 | ** |
| Log of real equivalent income | 9.373 | 0.002 | 9.511 | 0.008 | *** | 9.448 | 0.003 | 9.588 | 0.007 | *** |
| Married | 0.538 | 0.002 | 0.601 | 0.006 | *** | 0.548 | 0.002 | 0.528 | 0.006 | *** |
| Cohabiting | 0.122 | 0.001 | 0.148 | 0.005 | *** | 0.126 | 0.001 | 0.157 | 0.004 | *** |
| Divorce | 0.071 | 0.001 | 0.058 | 0.003 | *** | 0.038 | 0.001 | 0.038 | 0.002 | |
| Separated | 0.024 | 0.000 | 0.018 | 0.002 | *** | 0.014 | 0.000 | 0.010 | 0.001 | *** |
| Widowed | 0.028 | 0.001 | 0.021 | 0.002 | *** | 0.007 | 0.000 | 0.007 | 0.001 | |
| Health = 2 | 0.084 | 0.001 | 0.083 | 0.004 | | 0.066 | 0.001 | 0.070 | 0.003 | |
| Health = 3 | 0.212 | 0.001 | 0.228 | 0.005 | *** | 0.193 | 0.001 | 0.210 | 0.005 | *** |
| Health = 4 | 0.446 | 0.002 | 0.465 | 0.006 | *** | 0.445 | 0.002 | 0.448 | 0.006 | |
| Health = 5 | 0.235 | 0.001 | 0.198 | 0.005 | *** | 0.277 | 0.002 | 0.253 | 0.005 | *** |
| Number of days spent in hospital in t-1 | 0.825 | 0.018 | 0.701 | 0.066 | | 0.574 | 0.019 | 0.407 | 0.034 | *** |
| Own home outright | 0.725 | 0.001 | 0.760 | 0.006 | *** | 0.750 | 0.002 | 0.778 | 0.005 | *** |
| First degree | 0.105 | 0.001 | 0.073 | 0.003 | *** | 0.108 | 0.001 | 0.097 | 0.003 | *** |
| Higher degree | 0.022 | 0.000 | 0.020 | 0.002 | | 0.030 | 0.001 | 0.028 | 0.002 | |
| Vocational | 0.062 | 0.001 | 0.068 | 0.003 | * | 0.071 | 0.001 | 0.075 | 0.003 | |
| A-level | 0.182 | 0.001 | 0.187 | 0.005 | | 0.216 | 0.001 | 0.274 | 0.005 | *** |
| O-level | 0.292 | 0.001 | 0.343 | 0.006 | *** | 0.243 | 0.002 | 0.276 | 0.005 | *** |
| GCSE | 0.061 | 0.001 | 0.066 | 0.003 | | 0.056 | 0.001 | 0.060 | 0.003 | |
| Number of dependent children | 0.674 | 0.003 | 0.566 | 0.012 | *** | 0.586 | 0.003 | 0.508 | 0.010 | *** |

Notes: The number of observations per group is the following: Women (non-winners) = 94,796; Men (non-winners) = 81,090; Women (winners) = 5,929; Men (winners) = 8,005. T-test represents a balance test for i) women (non-winners) vs. women (winners), and ii) men (non-winners) vs. men (winners), with the null hypothesis = the two means are the same. ***<1%; **<5%; *<1%.

## Table 2A: Lottery gains and individual characteristics – Exogeneity test

| | Real lottery gains in t | | |
|---|---|---|---|
| | *Men* | *Women* | *All* |
| Female | | | -168.4* |
| | | | (96.07) |
| Self-employed in t-1 | -129.3 | 47.31 | -62.07 |
| | (103.9) | (76.63) | (66.33) |
| Age in t-1 | 32.60*** | -0.493 | 24.28*** |
| | (9.676) | (10.26) | (9.187) |
| Age squared in t-1 | -0.381*** | 0.00527 | -0.283*** |
| | (0.109) | (0.0998) | (0.105) |
| Log of real equivalent income in t-1 | -11.07 | 67.99** | 1.684 |
| | (55.09) | (34.18) | (34.12) |
| Married in t-1 | -100.6 | 18.91 | 16.94 |
| | (95.42) | (43.65) | (74.57) |
| Cohabiting in t-1 | -4.347 | 43.54 | 77.54 |
| | (42.79) | (71.86) | (63.88) |
| Divorce in t-1 | -40.77 | 13.91 | -47.46 |
| | (56.81) | (63.66) | (53.26) |
| Separated in t-1 | -48.95 | -25.82 | -28.26 |
| | (58.35) | (71.81) | (52.01) |
| Widowed in t-1 | 519.8 | 78.40 | 173.1 |
| | (379.3) | (64.82) | (138.3) |
| Health = 2 in t-1 | 45.56 | 73.62 | 90.69* |
| | (48.00) | (63.94) | (49.52) |
| Health = 3 in t-1 | -12.64 | 36.86 | 3.221 |
| | (47.11) | (49.68) | (46.64) |
| Health = 4 in t-1 | -45.98 | -15.23 | -82.00 |
| | (47.53) | (51.65) | (60.91) |
| Health = 5 in t-1 | -46.63 | 131.9 | 40.50 |
| | (49.10) | (80.75) | (111.3) |
| Number of days spent in hospital in t-1 | -7.283** | -0.122 | -4.143* |
| | (3.625) | (1.889) | (2.292) |
| Own home outright in t-1 | -62.91 | -78.66** | -125.9* |
| | (44.32) | (31.43) | (69.85) |
| First degree in t-1 | -141.2* | -127.8** | -54.70 |
| | (83.80) | (49.97) | (52.35) |
| Higher degree in t-1 | -303.2*** | -123.3* | -160.5*** |
| | (88.25) | (63.60) | (43.79) |
| Vocational in t-1 | 54.03 | -69.47 | 301.1 |
| | (270.4) | (50.16) | (361.8) |
| A-level in t-1 | -106.1 | 51.91 | -37.34 |
| | (86.47) | (150.1) | (37.90) |
| O-level in t-1 | -36.85 | -44.94 | 36.21 |
| | (82.50) | (42.28) | (51.61) |
| GCSE in t-1 | 362.7 | -83.64* | 345.8 |
| | (476.8) | (46.53) | (421.0) |
| Number of dependent children in t-1 | 27.33 | 64.58 | 14.90 |
| | (22.62) | (54.53) | (21.80) |
| Regional dummies | Yes | Yes | Yes |
| Wave dummies | Yes | Yes | Yes |
| Observations | 7,179 | 5,424 | 12,603 |
| Number of individuals | 2,907 | 2,467 | 5,374 |

Notes: Sample of winners in year t. Standard errors clustered at the individual level. ***<1%; **<5%; *<1%.

## Table 3A: Self-employment and lottery wins: Logit with random effects regressions – Full results

| | Self-employment Probability in *t* | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Female | -2.515*** | -2.513*** | -2.518*** | -2.586*** | -2.353*** | -2.516*** |
| | (0.0778) | (0.0778) | (0.271) | (0.291) | (0.386) | (0.409) |
| Top 25% winner in period t | | 0.230** | 0.182 | 0.110 | | |
| | | (0.107) | (0.155) | (0.169) | | |
| Female × Top 25% winner in period t | | | | 0.259 | | |
| | | | | (0.389) | | |
| Top 25% winner in period t-1 | | | | | 0.661*** | 0.547*** |
| | | | | | (0.180) | (0.193) |
| Female × Top 25% winner in period t-1 | | | | | | 0.423 |
| | | | | | | (0.452) |
| Age | 0.465*** | 0.465*** | 0.558*** | 0.558*** | 0.535*** | 0.529*** |
| | (0.025) | (0.025) | (0.070) | (0.071) | (0.087) | (0.084) |
| Age-squared | -0.005*** | -0.005*** | -0.006*** | -0.006*** | -0.005*** | -0.005*** |
| | (0.000) | (0.000) | (0.001) | (0.001) | (0.001) | (0.001) |
| Log of real equivalent household income | -0.221*** | -0.222*** | -0.194 | -0.195 | -0.269** | -0.268** |
| | (0.0346) | (0.0346) | (0.123) | (0.123) | (0.134) | (0.132) |
| Married | 0.486*** | 0.487*** | 0.380 | 0.381 | -0.0537 | -0.0556 |
| | (0.123) | (0.123) | (0.385) | (0.385) | (0.421) | (0.408) |
| Cohabiting | 0.522*** | 0.522*** | 0.594 | 0.592 | 0.599 | 0.595 |
| | (0.114) | (0.114) | (0.417) | (0.417) | (0.429) | (0.418) |
| Divorced | -0.0228 | -0.0224 | -0.234 | -0.235 | 0.161 | 0.139 |
| | (0.175) | (0.175) | (0.566) | (0.566) | (0.564) | (0.553) |
| Separated | 0.143 | 0.148 | -0.671 | -0.676 | 0.674 | 0.679 |
| | (0.189) | (0.189) | (0.772) | (0.773) | (0.713) | (0.697) |
| Widowed | -0.325 | -0.325 | -1.495 | -1.486 | -0.256 | -0.298 |
| | (0.468) | (0.468) | (1.731) | (1.738) | (1.411) | (1.371) |
| Health: poor | 0.576*** | 0.577*** | 0.508 | 0.508 | 0.229 | 0.253 |
| | (0.158) | (0.158) | (0.473) | (0.475) | (0.588) | (0.583) |
| Health: fair | 0.940*** | 0.941*** | 0.876* | 0.878* | 0.340 | 0.362 |
| | (0.159) | (0.159) | (0.491) | (0.493) | (0.649) | (0.642) |
| Health: good | 1.146*** | 1.147*** | 1.046** | 1.051** | 0.534 | 0.557 |
| | (0.160) | (0.160) | (0.488) | (0.491) | (0.650) | (0.643) |
| Health: excellent | 1.195*** | 1.196*** | 1.018** | 1.019** | 0.991 | 1.015 |
| | (0.163) | (0.163) | (0.503) | (0.506) | (0.651) | (0.644) |
| Number of days spent in hospital last year | -0.016*** | -0.016*** | -0.048* | -0.048* | -0.036 | -0.034 |
| | (0.006) | (0.006) | (0.025) | (0.025) | (0.025) | (0.024) |
| Homeowner | 0.422*** | 0.421*** | 0.834*** | 0.836*** | 0.620* | 0.615** |
| | (0.0887) | (0.0887) | (0.280) | (0.280) | (0.323) | (0.312) |
| Qualification: 1st degree | 0.525*** | 0.525*** | 0.562 | 0.562 | 0.863* | 0.860** |
| | (0.137) | (0.137) | (0.393) | (0.393) | (0.442) | (0.431) |
| Qualification: higher degree | 0.270 | 0.270 | 0.109 | 0.112 | -0.115 | -0.115 |
| | (0.218) | (0.218) | (0.578) | (0.578) | (0.708) | (0.689) |
| Qualification: Vocational | 0.298* | 0.296* | 0.317 | 0.314 | 0.577 | 0.569 |
| | (0.170) | (0.170) | (0.425) | (0.425) | (0.475) | (0.463) |
| Qualification: A-level | 0.371*** | 0.369*** | 0.698** | 0.703** | 0.952*** | 0.948*** |
| | (0.118) | (0.118) | (0.311) | (0.310) | (0.355) | (0.345) |
| Qualification: O-level | 0.184* | 0.181 | 0.273 | 0.275 | 0.237 | 0.233 |
| | (0.111) | (0.111) | (0.302) | (0.302) | (0.367) | (0.356) |
| Qualification: GCSE | 0.246 | 0.245 | 0.559 | 0.566 | 0.526 | 0.526 |
| | (0.188) | (0.188) | (0.498) | (0.498) | (0.572) | (0.560) |
| Number of dependent children | 0.00659 | 0.00651 | 0.252** | 0.252** | 0.295** | 0.299** |
| | (0.0399) | (0.0399) | (0.109) | (0.109) | (0.121) | (0.119) |

**Marginal effects at the mean**

| | | | | | | |
|---|---|---|---|---|---|---|
| Female | -0.072*** | -0.072*** | -0.061*** | -0.061*** | -0.060*** | -0.062*** |
| | (0.002) | (0.002) | (0.006) | (0.006) | (0.00655) | (0.006) |
| Top 25% winner in period t | | 0.006** | 0.005 | | | |
| | | (0.003) | (0.004) | | | |
| Male: Top 25% winner in period t | | | | 0.004 | | |
| | | | | (0.006) | | |
| Female: Top 25% winner in period t | | | | 0.004 | | |
| | | | | (0.004) | | |
| Top 25% winner in period t-1 | | | | | 0.018*** | |
| | | | | | (0.005) | |
| Male: Top 25% winner in period t-1 | | | | | | 0.022*** |
| | | | | | | (0.008) |
| Female: Top 25% winner in period t-1 | | | | | | 0.014** |
| | | | | | | (0.007) |
| Control variables | Yes | Yes | Yes | Yes | Yes | Yes |
| Regional dummies | Yes | Yes | Yes | Yes | Yes | Yes |
| Wave dummies | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 189,820 | 189,820 | 13,934 | 13,934 | 11,952 | 11,952 |

Notes: The dependent variable is a dummy variable that takes the value one if the respondent is self-employed in period *t* and zero otherwise. The figures in Columns 1 and 2 refer to the working-age respondents while figures in the remaining Columns refer to the lottery winners in the year of winning. The average real winning for the top 25% winners = £831.16 (minimum = £85.66; maximum = £184,672.20), and the bottom 75% winners = £25.69 (minimum = £1.64; maximum = £85.23). Control variables in each regression are the age, the age squared, the log of real equivalent household income, dummies for marital status, dummies for self-reported health status, education dummies, homeownership, the number of days spent in hospital in year t-1 and the number of dependent children. Standard errors are clustered at the individual level. *<10%; **<5%; ***<1%.

**Table 4A: Linear probability model of self-employment: Generalised least squares with random effects regression**

| | Self-employment Probability in *t* | |
|---|---|---|
| | (1) | (2) |
| Female | -0.073*** | -0.073*** |
| | (0.00688) | (0.00704) |
| Top 25% winner in period t-1 | 0.018*** | 0.018** |
| | (0.005) | (0.007) |
| Female × Top 25% winner in period t-1 | | 0.000 |
| | | (0.010) |
| Observations | 12,018 | 12,018 |
| Number of individuals | 5,361 | 5,361 |

Notes: The dependent variable is a dummy variable that takes the value one if the respondent is self-employed in period *t* and zero otherwise. The figures refer to the lottery winners the year of winning. The average real winning for the top 25% winners = £831.16 (minimum = £85.66; maximum = £184,672.20), and the bottom 75% winners = £25.69 (minimum = £1.64; maximum = £85.23). Control variables in each regression are the age, the age squared, the log of real equivalent household income, dummies for marital status, dummies for self-reported health status, education dummies, homeownership, the number of days spent in hospital in year *t-1* and the number of dependent children. The regressions also include region and wave dummies. Standard errors are clustered at the individual level. *<10%; **<5%; ***<1%.

**Table 5A: Self-employment and lottery wins accounting for risk-aversion: Logit with random effects regressions**

| | Self-employment Probability in *t* |
|---|---|
| Female | -2.551*** |
| | (0.452) |
| Top 25% winner in period t-1 | 0.397* |
| | (0.229) |
| Female × Top 25% winner in period t-1 | 0.469 |
| | (0.508) |
| Risk Aversion | 0.425*** |
| | (0.0780) |
| **Marginal effects at the mean** | |
| Female | -0.058*** |
| | (0.008) |
| Male: Top 25% winner in period t-1 | 0.016* |
| | (0.009) |
| Female: Top 25% winner in period t-1 | 0.011 |
| | (0.007) |
| Risk Aversion | 0.011*** |
| | (0.002) |
| Observations | 8,895 |
| Number of individuals | 3,568 |

Notes: The dependent variable is a dummy variable that takes the value one if the respondent is self-employed in period *t* and zero otherwise. The figures refer to the lottery winners the year of winning. The average real winning for the top 25% winners = £831.16 (minimum = £85.66; maximum = £184,672.20), and the bottom 75% winners = £25.69 (minimum = £1.64; maximum = £85.23). Control variables in each regression are all measured in *t-1* and are the age, the age squared, the log of real equivalent household income, dummies for marital status, dummies for self-reported health status, education dummies, homeownership, the number of days spent in hospital in year *t-1* and the number of dependent children. The regressions also include region and wave dummies. Standard errors are clustered at the individual level. *<10%; **<5%; ***<1%.

**Table 6A: Self-employment and lottery wins: Logit with random effects regressions and lottery wins categories**

| | Self-employment Probability in *t* | |
|---|---|---|
| | (1) | (2) |
| Female | -2.778*** | -2.263*** |
| | (0.660) | (0.492) |
| Log (real lottery win) in period t-1 | 0.176*** | |
| | (0.060) | |
| Female × Log (real lottery win) in period t-1 | 0.110 | |
| | (0.141) | |
| Q2: Log(real lottery win) in period t-1 | | 0.244 |
| | | (0.214) |
| Q3: Log(real lottery win) in period t-1 | | 0.177 |
| | | (0.216) |
| Q4: Log(real lottery win) in period t-1 | | 0.706*** |
| | | (0.235) |
| Female × Q2: Log(real lottery win) in period t-1 | | -0.507 |
| | | (0.487) |
| Female × Q3: Log(real lottery win) in period t-1 | | -0.264 |
| | | (0.482) |
| Female × Q4: Log(real lottery win) in period t-1 | | 0.160 |
| | | (0.529) |
| **Marginal effects at the mean** | | |
| Female | -0.062*** | -0.062*** |
| | (0.006) | (0.006) |
| Male: Log (real lottery win) in period t-1 | 0.007*** | |
| | (0.002) | |
| Female: Log (real lottery win) in period t-1 | 0.004** | |
| | (0.001) | |
| Male: Q2 of Log(real lottery win) in period t-1 | | 0.009 |
| | | (0.008) |
| Male: Q3 of Log(real lottery win) in period t-1 | | 0.007 |
| | | (0.008) |
| Male: Q4 of Log(real lottery win) in period t-1 | | 0.028*** |
| | | (0.009) |
| Female: Q2 of Log(real lottery win) in period t-1 | | -0.003 |
| | | (0.005) |
| Female: Q3 of Log(real lottery win) in period t-1 | | -0.001 |
| | | (0.005) |
| Female: Q4 of Log(real lottery win) in period t-1 | | 0.013* |
| | | (0.007) |
| Observations | 12,018 | 12,018 |
| Number of individuals | 5,361 | 5,361 |

Notes: The dependent variable is a dummy variable that takes the value one if the respondent is self-employed in period *t* and zero otherwise. The figures refer to the lottery winners a year after of winning who were not self-employed in *t-1*. Control variables in each regression are the age, the age squared, the log of real equivalent household income, dummies for marital status, dummies for self-reported health status, education dummies, homeownership, the number of days spent in hospital in year *t-1* and the number of dependent children. The regressions also include region and wave dummies. Standard errors are clustered at the individual level. *<10%; **<5%; ***<1%.

**Table 7A: Self-employment and lottery wins with lag control variables: Logit with random effects regressions**

| | Self-employment Probability in *t* | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Female | -2.386*** | -2.507*** | -2.367*** |
| | (0.419) | (0.389) | (0.643) |
| Top 25% winner in period t-1 | | 0.553*** | 0.548*** |
| | | (0.191) | (0.204) |
| Female × Top 25% winner in period t-1 | | 0.386 | 0.409 |
| | | (0.436) | (0.497) |
| In full-time employment in period t-1 | | | 0.001 |
| | | | (0.066) |
| Not in the labour force in period t-1 | | | -1.294* |
| | | | (0.675) |
| Unemployed in period t-1 | | | -0.877 |
| | | | (0.627) |
| **Marginal effects at the mean** | | | |
| Female | -0.061*** | -0.059*** | -0.057*** |
| | (0.006) | (0.013) | (0.007) |
| Top 25% winner in period t-1 | 0.018*** | | |
| | (0.005) | | |
| Male: Top 25% winner in period t-1 | | 0.022*** | 0.021*** |
| | | (0.007) | (0.008) |
| Female: Top 25% winner in period t-1 | | 0.014** | 0.014* |
| | | (0.007) | (0.008) |
| Observations | 12,083 | 12,083 | 12,079 |
| Number of individuals | 5,384 | 5,384 | 5,382 |

Notes: The dependent variable is a dummy variable that takes the value one if the respondent is self-employed in period *t* and zero otherwise. The figures refer to the lottery winners the year of winning. The average real winning for the top 25% winners = £831.16 (minimum = £85.66; maximum = £184,672.20), and the bottom 75% winners = £25.69 (minimum = £1.64; maximum = £85.23). Control variables in each regression are all measured in *t-1* and are the age, the age squared, the log of real equivalent household income, dummies for marital status, dummies for self-reported health status, education dummies, homeownership, the number of days spent in hospital in year *t-1* and the number of dependent children. The regressions also include region and wave dummies. Standard errors are clustered at the individual level. *<10%; **<5%; ***<1%

## Table 8A: Self-employment and lottery wins by gender: Logit with random effects regressions

| | Self-employment Probability in *t* | | | |
|---|---|---|---|---|
| | *Men* | *Women* | *Men* | *Women* |
| Top 25% winner in period t-1 | 0.580*** | 0.981*** | 0.592*** | 0.908*** |
| | (0.186) | (0.356) | (0.188) | (0.346) |
| **Marginal effects at the mean** | | | | |
| Top 25% winner in period t-1 | 0.023*** | 0.014** | 0.023*** | 0.013** |
| | (0.007) | (0.006) | (0.007) | (0.005) |
| Control variables in period t | Yes | Yes | No | No |
| Control variables in period t-1 | No | No | Yes | Yes |
| Regional dummies | Yes | Yes | Yes | Yes |
| Wave dummies | Yes | Yes | Yes | Yes |
| Observations | 6,822 | 5,025 | 6,881 | 5,096 |
| Number of individuals | 2,875 | 2,411 | 2,895 | 2,444 |

Notes: The dependent variable is a dummy variable that takes the value one if the respondent is self-employed in period *t* and zero otherwise. The figures refer to the lottery winners the year of winning. The figures refer to the lottery winners the year of winning. The average real winning for the top 25% winners = £831.16 (minimum = £85.66; maximum = £184,672.20), and the bottom 75% winners = £25.69 (minimum = £1.64; maximum = £85.23). Control variables in each regression are the age, the age squared, the log of real equivalent household income, dummies for marital status, dummies for self-reported health status, education dummies, homeownership, the number of days spent in hospital in year *t-1* and the number of dependent children. The regressions also include region and wave dummies. Standard errors are clustered at the individual level. *<10%; **<5%; ***<1%.

**Table 9A: The effect of longer lag lottery wins on self-employment: Logit with random effects regressions**

| | Self-employment Probability in *t* | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Female | -2.230*** | -2.141 | -2.322*** | -2.285*** |
| | (0.492) | (7.280) | (0.309) | (0.335) |
| Top 25% winner in period t-2 | 0.718*** | 0.682 | | |
| | (0.197) | (0.502) | | |
| Female × Top 25% winner in period t-2 | | 0.0728 | | |
| | | (1.716) | | |
| Top 25% winner in period t-3 | | | 0.584*** | 0.625*** |
| | | | (0.191) | (0.214) |
| Female × Top 25% winner in period t-3 | | | | -0.151 |
| | | | | (0.459) |
| **Marginal effects at the mean** | | | | |
| Female | -0.062*** | -0.062*** | -0.070*** | -0.069*** |
| | (0.007) | (0.009) | (0.007) | (0.007) |
| Top 25% winner in period t-2 | 0.022*** | | | |
| | (0.006) | | | |
| Male: Top 25% winner in period t-2 | | 0.025 | | |
| | | (0.017) | | |
| Female: Top 25% winner in period t-2 | | 0.011 | | |
| | | (0.054) | | |
| Top 25% winner in period t-3 | | | 0.019*** | |
| | | | (0.006) | |
| Male: Top 25% winner in period t-3 | | | | 0.026*** |
| | | | | (0.008) |
| Female: Top 25% winner in period t-3 | | | | 0.007 |
| | | | | (0.006) |
| Observations | 10,345 | 10,345 | 8,914 | 8,914 |
| Number of individuals | 4,719 | 4,719 | 4,186 | 4,186 |

Notes: The dependent variable is a dummy variable that takes the value one if the respondent is self-employed in period *t* and zero otherwise. The figures refer to the lottery winners the year of winning. The average real winning for the top 25% winners = £831.16 (minimum = £85.66; maximum = £184,672.20), and the bottom 75% winners = £25.69 (minimum = £1.64; maximum = £85.23). Control variables in each regression are the age, the age squared, the log of real equivalent household income, dummies for marital status, dummies for self-reported health status, education dummies, homeownership, the number of days spent in hospital in year *t-1* and the number of dependent children. The regressions also include region and wave dummies. Standard errors are clustered at the individual level. *<10%; **<5%; ***<1%.